THE
PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 98, NO. 2 APRIL 15, 1955

Partition Function for a System of Interacting Bose-Einstein Particles*†

S. T. BUTLER, *School of Physics, The University of Sydney, Sydney, Australia*

AND

M. H. FRIEDMAN, *Pupin Physics Laboratory, Columbia University, New York, New York*

(Received January 6, 1955)

It is shown by a rigorous development from first principles that a relatively simple form for the partition function of an imperfect Bose-Einstein gas can be obtained. This form is that of a $3N$-dimensional integral (N being the total number of particles) the integrand of which can be completely determined for any given interparticle interaction provided certain, in general small, many-body quantum mechanical effects are ignored. The final evaluation of the partition function in closed form requires that this integrand be approximated by functions sufficiently tractable that the $3N$-dimensional integral can be performed.

I. INTRODUCTION

THE derivation of the thermodynamic properties of a system of ideal (noninteracting) Bose-Einstein particles is well known.[1-3] Treatment of the physical case, however, where interparticle interactions are included, is in general much more difficult and the partition function for such a system has as yet been accurately evaluated only under very special circumstances, e.g., for temperatures sufficiently high that an expansion of the interaction terms in inverse powers of the temperature is rapidly convergent, or for very low densities (virial coefficient expansion).

In this paper we will show, by a rigorous development from first principles, that a relatively quite simple form for the partition function can be obtained. This form is that of a $3N$-dimensional integral (N being the total number of particles) the integrand of which can be completely determined for any given interparticle interaction provided certain, in general small, many-body quantum mechanical effects are ignored. The final evaluation of the partition function in closed form then requires that this integrand be approximated by functions sufficiently tractable that the $3N$-dimensional integration can be performed.

It is seen that the above-mentioned many-body effects contribute only to the long-range density fluctuations discussed, e.g., by Feynman,[4] and are therefore important only for temperatures sufficiently low so that the average particle wavelength is greater than the average interparticle separation. We hope to give a more detailed discussion of these effects in a subsequent publication.

A discussion of the application of these results to the case of He^4 is given in the last section, and a preliminary rather crude model, similar to that recently proposed by Feynman,[5] which approximates to some of the more important features of the partition function in this case, is set up. The results of this model are evaluated in the succeeding paper (hereafter called II).

II. DEFINITIONS OF THE PARTITION FUNCTION

The partition function to be evaluated is defined by

$$Z=\sum_j \exp(-\beta U_j), \tag{1}$$

where $\beta=1/kT$, and U_j are the energy levels of the system. This may, alternatively, be written as the Slater sum:

$$Z=\sum_\alpha(\Psi_\alpha, \exp-\{\beta(H_0+H')\}\Psi_\alpha), \tag{2}$$

* Supported in part by the Office of Naval Research.

† This work was commenced when the authors were associated at the Laboratory of Nuclear Studies, Cornell University, Ithaca, New York. One of us (M.H.F.) then held a National Science Foundation Post-Doctoral Fellowship. A preliminary report of this work was given at the 1953 Washington Meeting of the American Physical Society [Phys. Rev. 91, 465 (1953)].

[1] A. Einstein, Ber. Berl. Akad. 261 (1924); 3 (1925).

[2] F. London, Phys. Rev. 54, 947 (1938).

[3] B. Kahn and G. E. Uhlenbeck, Physica 5, 399 (1938).

[4] R. P. Feynman, Phys. Rev. 91, 1301 (1953).

[5] R. P. Feynman, Phys. Rev. 91, 1291 (1953).

where H_0 is the kinetic energy operator for the N-particles, H' is the interaction energy of the complete system, and the Ψ_α are any *appropriately symmetrized* wave functions which form a complete and orthonormal set over the volume of the container of the system.

Equation (2) is, of course, quite general irrespective of the type of statistics holding. The wave functions Ψ_α must merely be chosen in accord with whatever statistics are being dealt with. In this paper we will be concerned exclusively with Bose-Einstein statistics. For this case it is easy to show that (2) takes the form

$$Z=(1/N!)\sum_P\sum_j(\Phi_j{}^P,\ \exp\{-\beta(H_0+H')\}\Phi_j), \quad (3)$$

where N is the number of particles in the system, and the Φ_j are any complete and normalized set of *unsymmetrized* (Boltzmann) wave functions. The symbol P denotes some permutation of the set of coordinates $\mathbf{r}_1, \mathbf{r}_2, \cdots \mathbf{r}_N$ of the particles, and $\sum_P$ implies a summation over all permutations. Thus $\Phi_j{}^P$ is the same as Φ_j except that a certain interchange of coordinates is effected by the P. Because of the summation on P any set of wave functions Φ_j for a Boltzmann system is permissible—it is irrelevant in which way the wave functions describing states of the same energy, for example, are constructed.

Equation (3) forms our starting point for evaluation of Z.

III. EVALUATION OF THE PARTITION FUNCTION

In this section, we discuss the question of the evaluation of (3). We will first consider very briefly the case of the ideal gas, since the form then taken by (3) will be interesting for the purpose of comparison when we come to interacting systems.

1. Perfect Gas

Here we may write

$$\Phi_j=\frac{1}{(V)^{N/2}}\exp\left(\frac{i}{\hbar}\sum_{k=1}^{N}\mathbf{p}_k\cdot\mathbf{r}_k\right), \quad (4)$$

$$\Phi_j=\frac{1}{(V)^{N/2}}\exp\left(\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}\right), \quad (4a)$$

the state j now referring to a particular set of the momenta $\mathbf{p}_1, \cdots \mathbf{p}_N$. In (4a) we introduce, for convenience of notation, the $3N$-dimensional vectors $\mathbf{p}$ and $\mathbf{r}$.

If we insert this in (3), and replace the summation over Boltzmann states j by an integration over momenta, i.e.,

$$\sum_j\rightarrow\frac{V^N}{(2\pi\hbar)^{3N}}\int d\mathbf{p},$$

we find, after integrating over the momenta, that

$$Z=\frac{1}{N!}\frac{1}{\lambda^{3N}}\sum_P\int_V d\mathbf{r}\exp\left\{-\frac{\pi}{\lambda^2}(\mathbf{r}-\mathbf{r}_P)^2\right\}. \quad (5)$$

Here

$$\lambda^2=2\pi\hbar^2\beta/m, \quad (6)$$

and $\mathbf{r}_P$ is that position vector which replaces $\mathbf{r}$ in the permutation P.

This is the partition function as obtained by Kahn and Uhlenbeck[5] and yields the well-known properties of the perfect gas, including the Bose-Einstein condensation. The replacement of the summation over the states of the Boltzmann system by an integration is a legitimate procedure as long as the subsequent integration over coordinates is correctly limited by the volume of the container.[6]

2. Interacting Particles

Because of the simplicity of the wave functions (4) it is tempting to try to evaluate (3) in the general case also by continuing to choose the Φ_j to be plane waves. The difficulty then is clearly the evaluation of $\exp\times\{-\beta(H_0+H')\}\Phi_j$.

Two methods which have been developed to tackle this problem may be summarized as follows: Let us introduce an operator $S(\beta)$ such that

$$\exp\{-\beta(H_0+H')\}=S(\beta)\exp(-\beta H_0). \quad (7)$$

Then the operation of $\exp(-\beta H_0)$ on Φ_j is trivial [assuming the Φ_j to be plane waves—Eq. (4)]. In order to evaluate $S(\beta)$, we note that this operator must satisfy the equation:

$$\partial S/\partial\beta=-S(\beta)H'(\beta), \quad (8)$$

where

$$H'(\beta)=\exp(-\beta H_0)H'\exp(+\beta H_0).$$

Equation (8) is subject, of course, to the boundary condition $S\rightarrow 1$ as $\beta\rightarrow 0$.

Equation (8) may now be solved by an iteration procedure in a manner familiar from electrodynamics and yields for S the ordered exponential:

$$S(\beta)=1+\sum_{n=1}^{\infty}(-1)^n\int_0^\beta d\beta_1\int_0^{\beta_1}d\beta_2\cdots\int_0^{\beta_{n-1}}\times d\beta_n H'(\beta_n)\cdots H'(\beta_1). \quad (9)$$

Moreover, it is possible to carry out the operation of each term of (9) on the plane-wave Φ_j's.

This expression, derived by Goldberger and Adams,[7] is of use however only at temperatures sufficiently high that the interaction involved in H' can be considered as weak. It is essentially an expansion in powers of the ratio of the potential to kinetic energies. But in most physical systems of interest the interaction between

[6] Kahn and Uhlenbeck (reference 3) actually let these limits be infinite, but devise a prescription for obtaining the partition function below the condensation point. When the true limits of integration are retained, however, the evaluation of the integral (5) yields a form for the partition function identical to that obtained by London (reference 2), and the thermodynamic properties both above and below the transition point can immediately be derived (see Appendix, paper II).

[7] M. L. Goldberger and E. N. Adams, J. Chem. Phys. **20**, 240 (1952).

particles is extremely strong (e.g., very strongly repulsive at short distances) and it is hopeless to attempt to employ (9) at anything but high temperatures.

Another approach is to note that, instead of (8), an alternative equation for S is

$$\partial S/\partial\beta=-H'S(\beta)+[S(\beta),H_0], \tag{10}$$

where [ ] signifies that the commutator be taken. If now (10) be solved by iteration, the first step being to ignore the commutator of S and H_0 on the right-hand side, we obtain

$$S(\beta)=\exp(-\beta H')\Big\{1+\int_0^\beta d\beta_1 \times\{H_0-\exp(+\beta_1 H')H_0\exp(-\beta_1 H')\}+\cdots\Big\}. \tag{11}$$

Here we have derived the semiclassical expansion of Wigner[8]—as extended by Mayer and Band.[9] The consequences of the first term of (11) have, of course, been studied by many workers, particularly as applied to He^4 (e.g., Matsubara[10]). As we will see later, this term gives Z accurately for temperatures sufficiently high that $\lambda\ll r_0$, where r_0 is the range of the interparticle interaction. In the case of He^4, for example, this means temperatures $>80°K$.

The evaluation of higher terms in (11), operating on the plane-wave Φ_j's, can in principle be carried out. However, they become prohibitively complicated after the first two or three. But a very large number of terms of (11) become important at the low temperatures of so much physical interest (e.g., the vicinity of the λ-transition in He^4). The expansion (11) cannot be employed therefore for a study of this low-temperature region.

The method which we will now develop differs radically from the above approaches. Rather than expanding in an infinite series and being troubled by convergence problems, we will deal with an infinite product which may be employed for evaluation of the partition function not only for high temperatures but also for low.

Let us write

$$\exp\{-\beta(H_0+H')\}=\lim_{n\to\infty}[\exp(-\beta H'/n)\exp(-\beta H_0/n)]^n =\lim_{n\to\infty}\exp(-\beta H'/n)\exp(-\beta H_0/n) \times\exp(-\beta H'/n)\cdots\exp(-\beta H_0/n). \tag{12}$$

We also make use of the theorem, established in the Appendix, that (we retain the $3N$-dimensional vector notation):

$$\exp(-\beta H_0/n)f(\mathbf{r})=\frac{1}{{\lambda_n}^{3N}}\int d\boldsymbol{\delta}\exp\left(-\frac{\pi}{{\lambda_n}^2}\delta^2\right) \times f(\mathbf{r}-\boldsymbol{\delta}), \tag{13}$$

where ${\lambda_n}^2=\lambda^2/n$. Thus the effect of the operator $\exp(-\beta H_0/n)$ is merely to "smear" or average the function $f(\mathbf{r})$ over small displacements of each coordinate by amounts $\lesssim\lambda_n$.

On substituting (12) in (3), making use of (13), and performing the integration over momenta exactly as in the case of the perfect gas, we obtain for the partition function:

$$Z=\lim_{n\to\infty}\frac{1}{N!}\frac{1}{\{{\lambda_n}^{3N}\}^n}\sum_P\int_V d\mathbf{r}d\boldsymbol{\delta}^{(1)}d\boldsymbol{\delta}^{(2)}\cdots d\boldsymbol{\delta}^{(n-1)} \times\exp\left\{-\frac{\pi}{{\lambda_n}^2}(\mathbf{r}-\mathbf{r}_P-\boldsymbol{\delta}^{(1)}-\boldsymbol{\delta}^{(2)}\cdots\boldsymbol{\delta}^{(n-1)})\right\}^2 \times\exp\left\{-\frac{\pi}{{\lambda_n}^2}\sum_{\nu=1}^{n-1}(\delta^{(\nu)})^2\right\}V_n(\mathbf{r})V_n(\mathbf{r}-\boldsymbol{\delta}^{(1)}) \times V_n(\mathbf{r}-\boldsymbol{\delta}^{(1)}-\boldsymbol{\delta}^{(2)})\cdots \times V_n(\mathbf{r}-\boldsymbol{\delta}^{(1)}-\boldsymbol{\delta}^{(2)}\cdots\boldsymbol{\delta}^{(n-1)}). \tag{14}$$

Here we have let

$$V_n(\mathbf{r})=\prod_{i>j}\exp[-\beta\mathfrak{V}(r_{ij})/n],$$

$\mathfrak{V}$ denoting the interaction between any two particles.

By completing the squares on the $\boldsymbol{\delta}$'s in the exponential of (14) and transforming to the resulting variables of integration, this equation now becomes

$$Z=\frac{1}{N!}\frac{1}{\lambda^{3N}}\sum_P\int_V d\mathbf{r}\exp\left\{-\frac{\pi}{\lambda^2}(\mathbf{r}-\mathbf{r}_P)\right\}^2 Q(\mathbf{r},\mathbf{r}_P), \tag{15}$$

where

$$Q(\mathbf{r},\mathbf{r}_P)=\lim_{n\to\infty}\frac{1}{\{{\lambda_n}^{3N}\}^{(n-1)}}\int d\boldsymbol{\delta}^{(1)}\cdots \times d\boldsymbol{\delta}^{(n-1)}\exp\left\{-\frac{\pi}{{\lambda_n}^2}\sum_{\nu=1}^{n-1}(\delta^{(\nu)})^2\right\}V_n\Big\{\mathbf{r}-A_1(\mathbf{r}-\mathbf{r}_P) -(1/2^{\frac{1}{2}})\boldsymbol{\delta}^{(1)}-\tfrac{1}{2}(\tfrac{2}{3})^{\frac{1}{2}}\boldsymbol{\delta}^{(2)}\cdots-\tfrac{1}{3}(\tfrac{3}{4})^{\frac{1}{2}}\boldsymbol{\delta}^{(3)} -\frac{1}{(n-1)}\left(\frac{n-1}{n}\right)^{\frac{1}{2}}\boldsymbol{\delta}^{(n-1)}\Big\}V_n\Big\{\mathbf{r}-A_2(\mathbf{r}-\mathbf{r}_P) -(\tfrac{2}{3})^{\frac{1}{2}}\boldsymbol{\delta}^{(2)}-\tfrac{2}{3}(\tfrac{3}{4})^{\frac{1}{2}}\boldsymbol{\delta}^{(3)}\cdots-\frac{2}{(n-1)}\left(\frac{n-1}{n}\right)^{\frac{1}{2}}\boldsymbol{\delta}^{(n-1)}\Big\} \cdots \times V_n\Big\{\mathbf{r}-A_{n-1}(\mathbf{r}-\mathbf{r}_P)-\left(\frac{n-1}{n}\right)^{\frac{1}{2}}\boldsymbol{\delta}^{(n-1)}\Big\}V_n(\mathbf{r}). \tag{16}$$

Here the coefficients A_m are given by the series:

$$A_m=m\sum_{\mu=0}^{n-1-m}\frac{1}{(\mu+m)(\mu+m+1)}=\left(1-\frac{m}{n}\right). \tag{17}$$

We see that in Eq. (15) we have an exact expression for Z which is of a form very similar to the first term of the semiclassical expansion—Eq. (11). Indeed the partion function (15) is the same as that of a "classical" Bose-Einstein gas, with an equivalent temperature-dependent interaction (which actually includes some many-body forces—see later). When n is taken to be merely unity, (15) is precisely the "classical" result.

[8] E. P. Wigner, Phys. Rev. **40**, 749 (1932).
[9] J. Mayer and W. Band, J. Chem. Phys. **15**, 191 (1947).
[10] T. Matsubara, Progr. Theoret. Phys. (Japan) **6**, 714 (1951).

Also in the limit of no interaction ($V_n=1$), we have $Q=1$ and (15) immediately passes into (5).

We shall now show that, on the basis of an approximation which is justified at all temperatures except those very close to absolute zero, (16) can be put in a form which enables Q to be evaluated for any given interparticle interaction.

Consider first, for simplicity, the evaluation of Q when $\mathbf{r}_P\equiv\mathbf{r}$ [the identity permutation $P=1$ (say)]. Let us go through the steps by which $Q_{P=1}$ would be calculated. The first operation is that of "smearing" the function $V_n(\mathbf{r})$, i.e., (we forego the $3N$-dimensional notation here) we have to perform

$$\frac{1}{\lambda_n^{3N}}\int d\boldsymbol{\delta}_1\cdots d\boldsymbol{\delta}_N\prod_\mu \exp\left(-\frac{\pi}{\lambda_n^2}\boldsymbol{\delta}_\mu^2\right)\times\prod_{i>j} v_n\left(\left|\mathbf{r}_{ij}-\frac{1}{\sqrt{2}}\boldsymbol{\delta}_{ij}\right|\right), \quad (18)$$

where

$$\mathbf{r}_{ij}=\mathbf{r}_i-\mathbf{r}_j, \quad \boldsymbol{\delta}_{ij}=\boldsymbol{\delta}_i-\boldsymbol{\delta}_j,$$

and

$$v_n(\mathbf{r}_{ij})=\exp\{-\beta\mathfrak{U}(\mathbf{r}_{ij})/n\}.$$

Now since we are interested in the limit $n\to\infty$ (i.e., $\lambda_n\to 0$), it is correct to expand the function v_n of (18) to second order in the δ's. On omitting terms which vanish on integration, (18) then becomes

$$\begin{aligned}&\frac{1}{\lambda_n^{3N}}\{\prod_{i>j} v_n(r_{ij})\}\int d\boldsymbol{\delta}_1\cdots d\boldsymbol{\delta}_N\left\{\prod_\mu \exp\left(-\frac{\pi}{\lambda_n^2}\boldsymbol{\delta}_\mu^2\right)\right\}\\ &\times\left[1+\tfrac{1}{2}\sum_{i>j}\frac{\delta_{ij}^2}{2!}\frac{\nabla^2 v_n(r_{ij})}{v_n(r_{ij})}\right.\\ &\left.+\tfrac{1}{2}\sum_{i>j>k}\frac{\{\boldsymbol{\delta}_{ij}\cdot\nabla v_n(r)_{ij}\}\boldsymbol{\delta}_{jk}\cdot\nabla v_n(r_{jk})}{v_n(r_{ij})v_n(r_{jk})}\right]\\ &=\{\prod_{i>j} v_n(r_{ij})\}\left[1+\frac{3}{4\pi}\lambda_n^2\sum_{i>j}\frac{\nabla^2 v_n(r_{ij})}{v_n(r_{ij})}\right.\\ &\left.-\frac{1}{4\pi}\lambda_n^2\sum_{i>j>k}\frac{\nabla v_n(r_{ij})\cdot\nabla v_n(r_{jk})}{v_n(r_{ij})v_n(r_{jk})}\right]. \quad (19)\end{aligned}$$

The second and third terms of (19) give the changes which have been affected by the operation $\exp(-\beta H_0/n)$ on the function $\prod v_n$. The third term represents three-body interactions, and shows that the final form for Q must involve many-body interactions of all orders up to N. However, on writing this term of (19) in the form

$$-\frac{\lambda_n^2}{4\pi}\{\prod_{i>j} v_n(r_{ij})\}\sum_{i>j>k}\frac{\mathbf{r}_{ij}\cdot\mathbf{r}_{jk}}{r_{ij}r_{jk}}\times\frac{\{[dv_n(r_{ij})/dr_{ij}]\times[dv_n(r_{jk})/dr_{jk}]\}}{v_n(r_{ij})v_n(r_{jk})},$$

we see that it should be small compared to the second (two-body) correction term of (19) provided we can assume the end result that the most probable distribution of particles in the gas[11] is one of uniformity throughout the container. In this event, those particles which are interacting with any given particle are on the average symmetrically disposed around this particle, and (14) would depend on fluctuations from this average for its contributions. It should therefore be much less important than the second term of (19), in which all terms of the summation are additive for the average distribution.

If we assume, therefore, that for physical systems of interest the third term of (19) is unimportant, we then find that, to second order in the expansion of v_n, (18) is identical to the expression

$$\frac{1}{(\lambda_n^3)^{N(N-1)/2}}\prod_{i>j}\left\{\int d\boldsymbol{\delta}\exp\left(-\frac{\pi}{\lambda_n^2}\delta^2\right)v_n(|\mathbf{r}_{ij}-\boldsymbol{\delta}|)\right. \quad (20)$$

Thus, on the foregoing approximation, the result of operation (18) is again a product of the form

$$\prod_{i>j} v_n'(r_{ij}).$$

In this manner we can work right through expression (16). We have, then, that

$$Q_{P=1}=\prod_{i>j} q(r_{ij}), \quad (21)$$

$$\begin{aligned}q(r_{ij})=\lim_{n\to\infty}&\frac{1}{\{\lambda_n^3\}^{(n-1)}}\int d\boldsymbol{\delta}_{ij}^{(1)}\cdots d\boldsymbol{\delta}_{ij}^{(n-1)}\exp\left\{-\frac{\pi}{\lambda_n^2}\sum_{\nu=1}^{n-1}(\delta_{ij}^{(\nu)})^2\right\}\\ &\times v_n\left\{\mathbf{r}_{ij}-\sqrt{2}\left[\frac{1}{\sqrt{2}}\boldsymbol{\delta}_{ij}^{(1)}+\tfrac{1}{2}(\tfrac{2}{3})^{\frac{1}{2}}\boldsymbol{\delta}_{ij}^{(2)}+\cdots\frac{1}{n-1}\left(\frac{n-1}{n}\right)^{\frac{1}{2}}\boldsymbol{\delta}_{ij}^{(n-1)}\right]\right\}\\ &\times v_n\left\{\mathbf{r}_{ij}-\sqrt{2}\left[(\tfrac{2}{3})^{\frac{1}{2}}\boldsymbol{\delta}_{ij}^{(2)}+\tfrac{2}{3}(\tfrac{3}{4})^{\frac{1}{2}}\boldsymbol{\delta}_{ij}^{(3)}+\cdots\frac{2}{n-1}\left(\frac{n-1}{n}\right)^{\frac{1}{2}}\boldsymbol{\delta}_{ij}^{(n-1)}\right]\right\}\\ &\cdots\\ &\times v_n\left\{\mathbf{r}_{ij}-\sqrt{2}\left[\left(\frac{\nu}{\nu+1}\right)^{\frac{1}{2}}\boldsymbol{\delta}_{ij}^{(\nu)}+\frac{\nu}{\nu+1}\left(\frac{\nu+1}{\nu+2}\right)^{\frac{1}{2}}\delta_{ij}^{(\nu+1)}+\cdots\frac{\nu}{n-1}\left(\frac{n-1}{n}\right)^{\frac{1}{2}}\boldsymbol{\delta}_{ij}^{(n-1)}\right]\right\}\\ &\cdots\\ &\times v_n\left\{\mathbf{r}_{ij}-\sqrt{2}\left(\frac{n-1}{n}\right)^{\frac{1}{2}}\boldsymbol{\delta}_{ij}^{(n-1)}\right\}v_n(\mathbf{r}_{ij}). \quad (22)\end{aligned}$$

[11] That is that distribution which contributes most to the integral of (15).

Although the limit of n going to infinity has to be taken to get $q(r_{ij})$ exactly, in practice it is possible to choose a sufficiently large but finite n such that q is obtained to a desired accuracy. Even so, the amount of numerical work necessary to evaluate $q(r_{ij})$ from (22) is, of course, extremely large. For this purpose, however, the following approximation to (22) is quite adequate:

$$q(r_{ij}) \simeq \lim_{n\to\infty} \frac{1}{(\lambda_n^3)^{(n-1)}} \int d\boldsymbol{\delta}_{ij}^{(1)} \cdots d\boldsymbol{\delta}_{ij}^{(n-1)} \exp\left\{-\frac{\pi}{\lambda_n^2}\sum_{\nu=1}^{n-1}(\delta_{ij}^{(\nu)})^2\right\}$$
$$\times v_n\left\{\mathbf{r}_{ij}-\sqrt{2}\left[\frac{1}{\sqrt{2}}\boldsymbol{\delta}_{ij}^{(1)}+(\tfrac{2}{3})^{\frac{1}{2}}\boldsymbol{\delta}_{ij}^{(2)}+\cdots\right]\right\}$$
$$\times v_n\{\mathbf{r}_{ij}-\sqrt{2}[(\tfrac{2}{3})^{\frac{1}{2}}\boldsymbol{\delta}_{ij}^{(2)}+(\tfrac{3}{4})^{\frac{1}{2}}\boldsymbol{\delta}_{ij}^{(3)}+\cdots]\}$$
$$\cdots$$
$$\times v_n\left\{\mathbf{r}_{ij}-\sqrt{2}\left[\left(\frac{\nu}{\nu+1}\right)^{\frac{1}{2}}\boldsymbol{\delta}_{ij}^{(\nu)}+\left(\frac{\nu+1}{\nu+2}\right)^{\frac{1}{2}}\boldsymbol{\delta}_{ij}^{(\nu+1)}+\cdots\right]\right\}$$
$$\cdots$$
$$\times v_n\left\{\mathbf{r}_{ij}-\sqrt{2}\left(\frac{n-1}{n}\right)^{\frac{1}{2}}\boldsymbol{\delta}_{ij}^{(n-1)}\right\} v_n(\mathbf{r}_{ij}). \quad (23)$$

The reason for this is that in each v_n of Eqs. (22) or (23) only the first few "smearing" variables $\boldsymbol{\delta}_{ij}^{(\nu)}$ are important, and (except for the very first v_n) it is largely irrelevant whether the smearing variables[12] are the square-bracketed terms of Eqs. (22) or (23).

Thus, from Eq. (23), q may be evaluated by performing $(n-1)$ operations of the form

$$\frac{1}{\lambda_n^3}\int d\boldsymbol{\delta} \exp\left\{-\frac{\pi}{\lambda_n^2}\delta^2\right\} f(|\mathbf{r}_{ij}-\alpha\boldsymbol{\delta}|). \quad (24)$$

For the first step the function f is to be merely v_n. The integral (24) (with $\alpha=1$) then yields a new function which is to be multiplied by v_n and again averaged by the integration of (24) [with α now $(4/3)^{\frac{1}{2}}$] and so on. After $(n-1)$ such steps have been performed, we multiply for the last time by $v_n(r_{ij})$ to get q.

For a given interaction and temperature, therefore, q can be obtained by performing n numerical integrations for each value of the argument r_{ij}. [The angle integrations of (24) are done analytically.] We note that, in computing q for a certain low temperature T/n, the various steps involved also give it for the temperatures T, $T/2$, $T/3\cdots T/n$.

A convenient way of obtaining an estimate of the accuracy involved by using a certain number of steps n is the following: instead of (12) we could equally well write:

$$\exp\{-\beta(H_0+H')\} = \lim_{n\to\infty}\left[\exp\left(-\frac{\beta}{n}H_0\right)\right.$$
$$\left.\times\exp\left(-\frac{\beta}{n}H'\right)\cdots\exp\left(-\frac{\beta}{n}H'\right)\right]. \quad (25)$$

Equation (25) gives precisely the same results as (12) except that in obtaining q one more smearing process [Eq. (22)] is performed after multiplying by the last v_n. Thus n should be sufficiently large that the difference between the two results obtained by carrying out the nth smearing and not doing so can be tolerated. Computations which we have performed indicate that the two curves thus obtained for q actually bracket the true curve, and of course tend closer to each other the larger n.

In the case of the general function Q, precisely the same considerations apply in expressing it in the form

$$Q=\prod_{i>j} q_P(\mathbf{r}_{ij},\mathbf{r}_{ij}^P), \quad (26)$$

where

$$q_P(\mathbf{r}_{ij},0)=q(r_{ij}).$$

Also, even when $\mathbf{r}_{ij}\neq\mathbf{r}_{ij}^P$, $q_P(\mathbf{r}_{ij},\mathbf{r}_{ij}^P)$ may be evaluated by going through the same steps described above for obtaining $q(r_{ij})$ from Eq. (23). Much more numerical work is now required, of course, since $(n-1)$ integrations have to be performed not only for various values of $\mathbf{r}_{ij}$, but also for various values of both $\mathbf{r}_{ij}^P$, and the angle between $\mathbf{r}_{ij}$ and $\mathbf{r}_{ij}^P$.[13] Once this has been carried

[12] Indeed, for all actual calculations which have been performed, the results for q obtained by taking only the first δ variable in each v_n differs very little from the results obtained from Eq. (23). For example, curves which are given in the next section for the case of He^4, obtained from Eq. (23) have the heights of their maxima above unity increased by less than 20 percent, when computed by the above much cruder approximation. The actual curve will lie between the two results, and should be much closer to the results of (23).

[13] This work is not necessary if $\lambda/\sqrt{\pi}$ is considerably less than the range r_0 of $\mathfrak{V}$; in this event the added arguments $A_m(\mathbf{r}_{ij}-\mathbf{r}_{ij}^P)$ will produce only a small smearing effect to $q(r_{ij})$ and may be ignored.

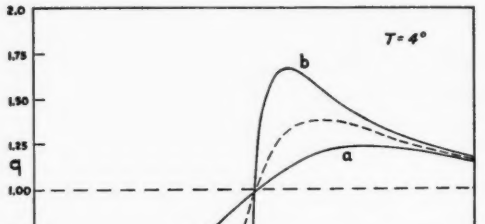


FIG. 1. Curves of $q(r_{ij})$ for $T=4°$, computed by taking $n=5$. Curve (b) is that obtained when the last smearing is not performed, and curve (a) is that obtained from (b) by performing the last smearing process. The dotted curve is a compromise between (a) and (b), estimating the actual behavior of q.

out for a given interaction $\mathfrak{V}$ however, the Q of (15) is completely known.

A more detailed discussion of the many-body quantum-mechanical effects which are ignored in the above will be made in a subsequent publication. It will be shown that they are associated with the long range density fluctuations, which are important only for temperatures such that $\lambda/2\sqrt{\pi} \gtrsim r_a$, where r_a is the average interparticle spacing.

III. APPLICATION TO He^4

In this section we give a preliminary discussion of the application of our results to the case of He^4 at low temperatures. We have made numerical calculations of the function $q(r_{ij})$ for this case, the interaction $\mathfrak{V}$ used being that of Slater and Kirkwood.[14,15] The magnitude

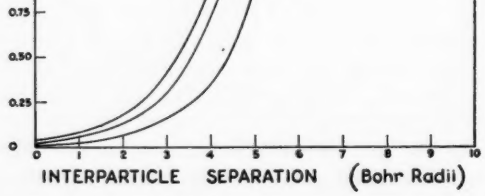


FIG. 2. Curves of $q(r_{ij})$ for a number of temperatures.

[14] J. C. Slater and J. G. Kirkwood, Phys. Rev. 37, 682 (1931).
[15] The precise behavior of $\mathfrak{V}$ for small interparticle distances (<2 Bohr radii) was found to be completely irrelevant to our results as long as it remained strongly repulsive.

of each of the steps used was such that β/n corresponded to a temperature of 20°. Thus, for example, 5 steps were taken for $T=4°$, 10 for $T=2°$, and 20 for $T=1°$.

In Fig. 1 we see the behavior thus found for $q(\mathbf{r}_{ij})$ at $T=4°$. Curve (b) is that obtained when the last smearing is not performed, and this passes into curve (a) by means of the last smearing process. The differences between the two curves show that, for a really precise determination of $q(r_{ij})$ at all points, we should have taken smaller intervals β/n. These curves are adequate for our present discussion, however, and we consider the actual behavior of $q(r_{ij})$ to be not far from a compromise between (a) and (b), such as the dotted curve.

Similar compromise curves for a number of other temperatures are given in Fig. 2.

It is interesting to see what a vast improvement is obtained by using the above rather small number of

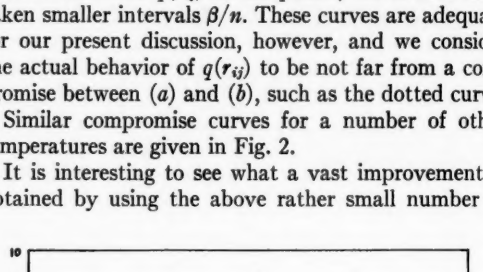


FIG. 3. The curves obtained for q at $T=4°$ when $n=1$, (a) being the unsmeared and (b) the smeared curve. The differences between these curves and those of Fig. 1 are to be noted.

steps over the crudest approximation of using only one step ($n=1$). At 4° the two curves for $n=1$ are shown in Fig. 3, (a) being the unsmeared and (b) the smeared curve. It is seen that the two are completely different from each other, and allow of no reasonable estimate of the actual behavior of $q(r_{ij})$. With increase of n from 1 to 5, however, these two curves pass into those of Fig. 1.

In this connection it is worthwhile to note how incorrect it is at these low temperatures to neglect all quantum-mechanical effects other than the statistics—i.e., to use only the first term of (11). This would give the result

$$q(r_{ij}) = \exp\{-\beta\mathfrak{V}(r_{ij})\}. \tag{27}$$

At 4° this is just the curve (a) of Fig. 3 which differs radically from the actual curve of Fig. 1. The differ-

ence is even more pronounced for lower temperatures. At 2° for example (i.e., the vicinity of the λ transition) Eq. (25) would give a peak due to the attractive portion[14] of $\mathfrak{V}$ of order of magnitude 100, instead of about 1.8 (Fig. 2).

The curves of Figs. 1 and 2 show clearly the characteristics of the function $q(r_{ij})$. Its most important property is that, at the low temperatures being considered, it is always very small for small interparticle separation. Thus the contribution to (15) from configurations in which two or more particles (whose coordinates are unaffected by the permutation P) are close together will be quite small.

It can easily be seen that the general function $q_P(\mathbf{r}_{ij},\mathbf{r}_{ij}{}^P)$ for $\mathbf{r}_{ij}{}^P \neq \mathbf{r}_{ij}$ also behaves in the above manner as a function of r_{ij}. The fact that for all the final steps of (16) the coefficients A_m [Eq. (17)] are very small ensures that $q_P(\mathbf{r}_{ij},\mathbf{r}_{ij}{}^P)$ has the same qualitative behavior as a function of r_{ij} as $q(r_{ij})$. We now have merely that the precise *small* value of q_P for small r_{ij} is dependent on $\mathbf{r}_{ij}{}^P$, and also that the precise height of the subsequent rise above unity is dependent on $\mathbf{r}_{ij}{}^P$.

A very simple model which approximates the above features is to ignore the difference between $q_P(\mathbf{r}_{ij},\mathbf{r}_{ij}{}^P)$ and $q(r_{ij})$, and to replace $q(r_{ij})$ by a step function which is zero for all r_{ij} up to some distance r_0, and unity thereafter.[16] The results of this model are evaluated in II, r_0 being taken to be $3.3a_0$ (where a_0 is the Bohr radius). It will be seen in II that (15) then yields a Bose-Einstein condensation point in quite good agreement with the λ-point of liquid helium. Moreover, the thermodynamic properties in the vicinity of the transition point are similar to those observed experimentally for liquid helium.

Of course in ignoring as it does the effect of the attractive part of $\mathfrak{V}$, exhibited in $q(r_{ij})$ as a rise above unity in the vicinity of $7a_0$, the above model does not yield the normal gas-liquid transition, and speaks of a gas rather than a liquid at the low temperatures. We consider it very probable that the remaining differences between the results of II and experiment are due to this neglect.

[16] Equation (15) is then very similar to the form for Z recently proposed by Feynman (see reference 5).

Since it is possible to numerically evaluate $q_P(\mathbf{r}_{ij},\mathbf{r}_{ij}{}^P)$ very accurately by the methods of Sec. III, however, it should be possible for the finer details to be calculated. What is required for this is that a function be found which simulates the behavior of $q_P(\mathbf{r}_{ij},\mathbf{r}_{ij}{}^P)$ in some detail but which is still sufficiently tractable for the integral of (15) to be evaluated. This question is being investigated.

The authors are indebted to Professor Bethe, Professor Dyson, Professor Placzek, and Professor Salpeter for discussions on the subject of this paper.

APPENDIX

We wish to prove Eq. (13), *viz.*,

$$\exp\left(-\frac{\beta}{n}H_0\right)f(\mathbf{r})=\frac{1}{\lambda_n{}^{3N}}\int d\boldsymbol{\delta}\exp\left(-\frac{\pi}{\lambda_n{}^2}\right)f(\mathbf{r}-\boldsymbol{\delta}).$$

Let

$$f(\mathbf{r})=\int d\mathbf{k}g(\mathbf{k})\exp(i\mathbf{k}\cdot\mathbf{r}).$$

Then

$$\exp\left(-\frac{\beta}{n}H_0\right)f(\mathbf{r})=\int d\mathbf{k}g(\mathbf{k})\exp(i\mathbf{k}\cdot\mathbf{r})\exp\left(-\frac{\lambda_n{}^2}{4\pi}k^2\right).$$

On introduction of the new variables $\mathbf{q}$ and $\mathbf{R}$ this may be written

$$\frac{1}{(2\pi)^{3N}}\int d\mathbf{R}d\mathbf{q}d\mathbf{k}g(\mathbf{k})\exp\left(-\frac{\lambda_n{}^2}{4\pi}q^2\right)\exp(i\mathbf{q}\cdot\mathbf{r})$$

$$\times\exp\{i\mathbf{R}\cdot(\mathbf{k}-\mathbf{q})\}=\frac{1}{\lambda_n{}^{3N}}\int d\mathbf{R}d\mathbf{k}g(\mathbf{k})\exp(i\mathbf{k}\cdot\mathbf{R})$$

$$\times\exp\left\{-\frac{\pi}{\lambda_n{}^2}(\mathbf{r}-\mathbf{R})^2\right\}=\frac{1}{\lambda_n{}^{3N}}\int d\mathbf{R}f(\mathbf{R})$$

$$\times\exp\left\{-\frac{\pi}{\lambda_n{}^2}(\mathbf{r}-\mathbf{R})^2\right\}.$$

On changing the variable of integration from $\mathbf{R}$ to $\boldsymbol{\delta}=\mathbf{r}-\mathbf{R}$, we immediately obtain Eq. (13).

Bose-Einstein Condensation of an Imperfect Gas*†

M. H. FRIEDMAN, *Pupin Physics Laboratory, Columbia University, New York, New York*

AND

S. T. BUTLER, *School of Physics, The University of Sydney, Sydney, Australia*

(Received January 6, 1955)

The form of the partition function for a system of interacting Bose-Einstein particles developed in the previous paper is treated approximately. A transition analogous to that occurring in the perfect gas is observed. The thermodynamic properties above and below this transition point are examined.

I. INTRODUCTION

IN the preceding paper,[1] the partition function for a system of Bose-Einstein particles has been shown to take on a relatively simple form for sufficiently low temperatures—(although not so low that the thermal wavelength exceeds the average interparticle distance). In this paper we shall develop a method whereby this partition function may be treated approximately, and will show that it then leads to a transition analogous to that for the perfect Bose-Einstein gas. We shall also be interested in examining the thermodynamic properties of this system in the neighborhood of the transition point, both above and below, comparing these results with those experimentally obtained for liquid helium.

II. REDUCTION OF THE PARTITION FUNCTION

For the sake of convenience we here rewrite the form of the partition function developed in reference 1.

$$Z=\frac{1}{N!}\frac{1}{\lambda^{3N}}\sum_P\int d\mathbf{r}_1\cdots d\mathbf{r}_N\exp\left\{-\frac{\pi}{\lambda^2}\sum_{i=1}^{N}|\mathbf{r}_j-\mathbf{r}_{Pj}|^2\right\}\times Q(\mathbf{r}_1\cdots\mathbf{r}_N), \quad (1)$$

where the function Q is to be taken as a product of step functions (one for each interparticle distance). If we were dealing with the ideal gas, Q would be identically unity everywhere. Under these circumstances, for a given permutation P, the Gaussians of the integrand of Eq. (1) breaks up into a product of factors, each factor involving a particular set of particles and no others.[2] The coordinates appearing in each factor may be diagrammatically thought of as being linked into a cycle by the terms $\exp[-(\pi/\lambda^2)|\mathbf{r}_j-\mathbf{r}_{Pj}|^2]$. Since the coordinates in each of these factors appear nowhere else, the integrand of Eq. (1) becomes a product of integrals, one for each factor. In reference 2 it is then noted that the resulting situation bears a close resemblance to that of Mayer's imperfect gas theory if one regards these integrals as analogous to his cluster integrals.

In the present case, however, there is an additional linking of the coordinates to one another through Q. Thus, although the Gaussian will still break up into a product of clusters (or cycles), the integral does not reduce to a product of integrals, and hence it becomes necessary to make some approximations as regards Q. We shall refer to the linking of the coordinates through the Gaussian factors as the λ linking, while that through the function Q as the Q linking.

Now, since $Q=\prod_{i>j}q(|\mathbf{r}_i-\mathbf{r}_j|)$ where $q(|\mathbf{r}_i-\mathbf{r}_j|)$ is a step function[1] of width r_0, its sole effect is to exclude regions of integration. We now note that $\lambda^2\gg r_0^2$ in the region of temperature of interest (namely, the λ-transition point of He^4). Hence, the Gaussians essentially remain constant over this excluded region. If we now restrict ourselves to the consideration of those circumstances in which the density is sufficiently low, so that on the average a particle can be moved between its neighbors and not be trapped in position, we may well approximate Q by

$$Q=\prod_{i>j}[1-v_0\Delta(r_{ij})]. \quad (2)$$

Here $v_0=(4\pi/3)r_0^3$ is the volume of the excluded region, and $\mathbf{r}_{ij}=\mathbf{r}_i-\mathbf{r}_j$. The function $\Delta(r_{ij})$ is similar to a Dirac delta function except that we replace $[v_0\Delta(r_{ij})]^n$ by $v_0\Delta(r_{ij})$ whenever it occurs. The above condition that the density be sufficiently low is the requirement that v_0 be less than or about $v/2$, where v is the average volume per particle, i.e., $v=V/N$.

We will now consider a particular permutation in the sum $\sum_P$ of Eq. (1). This then corresponds to the N particles being divided among some given number of clusters. We shall describe this number by saying that there are m_l clusters each containing l particles. We of course require $\sum_l lm_l=N$.

(a) Removal of the External Links

Let us now write the function Q as a product of two parts, one involving interactions between particles in the same cluster and one involving interactions be-

* Supported in part by the Office of Naval Research.

† This work was commenced when the authors were together at the Laboratory of Nuclear Studies, Cornell University, Ithaca, New York. One of us (M.H.F.) then held a National Science Foundation Post-Doctoral Fellowship. A preliminary report of this work was given at the 1953 Washington Meeting of the American Physical Society [Phys. Rev. **91**, 465 (1953)].

[1] S. T. Butler and M. H. Friedman, preceding paper [Phys. Rev. **98**, 287 (1955)].

[2] B. Kahn and G. E. Uhlenbeck, Physica **5**, 399 (1938).

tween particles of different clusters. We then have for the integral of Eq. (1):

$$S_P=\lim_{v_0'\to v_0} S_P'(v_0,v_0')$$

$$=\int d\mathbf{r}_1\cdots d\mathbf{r}_N \prod_i G_{li}Q_{li}(v_0)\prod_{\text{ext.}}[1-v_0'\Delta(r_{ij})]. \quad (3)$$

Here G_{li} represents the λ linking in a particular cluster i of size l, while $Q_{li}(v_0)$ represents the Q linking. The last product in the integrand of Eq. (3) denotes all the external Q linkings, and hence describes the interactions between the clusters. We now consider $\partial S_P'/\partial v_0'$. Thus

$$\partial S_P'/\partial v_0'=-\sum_{i>j}^{\text{ext.}}\int d\mathbf{r}_1\cdots d\mathbf{r}_N\Delta(r_{ij})\prod_i G_{li}Q_{li}\prod_{\text{ext.}} \times[1-v_0'\Delta(r_{pk})], \quad (4)$$

where the sum is to be taken over all pairs of particles not in the same cluster.

We would now like to show that the integral in Eq. (4) is to good approximations equal to S_P'/V.

Consider Eq. (3). Let us take two clusters, I and II (say), with particle m a member of the first and n the second. We choose a new set of coordinates of which one shall be $\mathbf{X}=\mathbf{r}_m-\mathbf{r}_n$; the other $(N-1)$ may be obtained by first letting one particle from each cluster transform into itself (except for cluster II), and taking the other $(l-1)$ coordinates for each cluster to be the relative coordinates between a particle and its preceding one in the ring. [There are of course l of these latter for each cluster, but we may only choose $(l-1)$ of them.] Let this set be called $\mathbf{y}_1,\cdots\mathbf{y}_{N-1}$. If we now neglect external Q linkings involving the particles in cluster II, with the exception of $q(\mathbf{X})=q(r_{mn})$, we have

$$S_P'\simeq\left(\int I_{mn}(\mathbf{y}_1,\cdots\mathbf{y}_{N-1})\right)\left(\int q(\mathbf{X})d\mathbf{X}\right)$$

$$=(V-v_0')\int I_{mn}(\mathbf{y}_1,\cdots\mathbf{y}_{N-1}), \quad (5)$$

where I_{mn} is some function independent of $\mathbf{X}$.

Similarly, going over to Eq. (4), we see that if particles i and j under the summation are taken to correspond to m and n of the preceding paragraph, and if in addition we neglect the same interactions that we did there, then the integral of Eq. (4) becomes,

$$\left(\int I_{mn}(\mathbf{y}_1,\cdots\mathbf{y}_{N-1})\right)\left(\int\Delta(X)d\mathbf{X}\right)$$

$$=\int I_{mn}(\mathbf{y}_1,\cdots\mathbf{y}_{N-1}), \quad (6)$$

which on comparison with Eq. (5) may be taken as $S_P'/(V-v_0')$, where the v_0' is of course negligible compared to V (since we shall eventually go to the limit $v_0'\to v_0$). Making this substitution, the sum in Eq. (4) may now be evaluated by multiplying by the total number of external links, so that

$$\partial S_P'/\partial v_0'=-\tfrac{1}{2}\{N(N-1)-\textstyle\sum_l m_l l(l-1)\}S_P'/V. \quad (7)$$

On solving (7) and taking the limit as $v_0'\to v_0$, we obtain

$$S_P=A_P\exp(-\tfrac{1}{2}\rho N v_0)\exp(\textstyle\sum_l m_l l^2 v_0/2V), \quad (8)$$

where ρ is the density of the particles. A_P is the constant of integration and is the integral of Eq. (3) with all external Q linkings omitted.

The error involved in the right-hand side of (7) due to the neglect of the external links of cluster II is of order one as compared to the total number of external links [i.e., the coefficient of S_P'/V on the right-hand side of (7)]. This is so, because the neglect leads to an error of order $\{l(N-l)/(\text{total number of external links})\}$ in each term of the summation of (4), where l is the size of cluster II. However,

$$\tfrac{1}{2}\textstyle\sum_l m_l l(N-l)=\tfrac{1}{2}\{N(N-1)-\sum_l m_l l(l-1)\}. \quad (9)$$

Thus the resulting Eq. (8) is actually quite accurate.

We will later demonstrate that to first order the term $\exp(m_l l^2 v_0/2V)$ in (8) is cancelled by the internal interactions. This is quite reasonable since the exponentials represent a "packing" effect due to containing the particles in a limited volume, and we might expect that interactions between particles in the same cluster should have as much of an effect on the packing as those between particles of different clusters. In other words, we expect the exponential term not to be dependent on l or m_l.

(b) Evaluation of the Internal Links

We now consider the factor A_P. Since all external Q linkings have been removed it breaks up into a product of integrals, one for each cluster. A typical one would be

$$VC_l=\int d\mathbf{r}_1\cdots d\mathbf{r}_l \times\exp\{-(\pi/\lambda^2)[r_{12}^2+\cdots r_{1l}^2]\}Q_l(\mathbf{r}_1,\cdots\mathbf{r}_l)\}. \quad (10)$$

For the present we shall forget that the limits of integration are really restricted to within a volume V and extend them out to infinity. In the following we shall be interested in obtaining an asymptotic expression for large l.

Performing $\partial C_l/\partial v_0$, we find

$$\frac{\partial C_l}{\partial v_0}=-\sum_{i>j}\frac{1}{V}\int d\mathbf{r}_1\cdots d\mathbf{r}_l\exp\left\{-\frac{\pi}{\lambda^2}[r_{12}^2+\cdots r_{1l}^2]\right\} \times\Delta(r_{ij})\prod_{k>l}'[1-v_0\Delta(r_{kl})]. \quad (11)$$

If in the integral we do the integration over $\mathbf{r}_i$, then as a result of $\Delta(r_{ij})$ we obtain a result that looks as if we had two clusters, one of size $(l-n)$ and one of size n joined at the point $\mathbf{r}_j$. If we ignore the Q linkings that occur between the two clusters, the result reduces to the product $C_{(l-n)}C_n$. At first glance one might object to this on the grounds that for $(l-n)\sim n$ the number of interactions neglected is of the same order as the number retained. The justification lies in examining the final result and noting that the major contribution to the right-hand side of Eq. (12) comes from terms with $n \ll l$, for which the neglect is valid.

Assuming for simplicity that l is odd, we have

$$\frac{\partial C}{\partial v_0} = -l\{C_{(l-1)} + \sum_{n=2}^{(l-1)/2} C_n C_{(l-n)}\}, \tag{12}$$

subject to the boundary condition that

$$C_l(v_0=0) = \lambda^{3(l-1)}/l^{\frac{5}{2}}. \tag{13}$$

We attempt a solution by assuming

$$C_l = (\lambda^3(1-\gamma v_0/\lambda^3)]^l/[l^{\frac{5}{2}}(1+\delta v_0/\lambda^3)\lambda^3], \tag{14}$$

where γ and δ are to be determined. We require that the form Eq. (14) satisfy Eq. (12), to order $1/l$. Substituting into the right-hand side and performing the sum to order $1/l^{5/2}$ (by means of the Euler expansion), we obtain

$$\frac{\partial C_l}{\partial v_0} = -lC_{(l-1)}\left\{1+\frac{(1-\gamma v_0/\lambda^3)}{(1+v_0/\lambda^3)}\frac{(l-1)^{\frac{1}{2}}}{l^{\frac{1}{2}}} \times\left[1.61-\frac{3.77}{l}\right]\right\}. \tag{15}$$

Substitution into the left-hand side yields

$$\frac{\partial C_l}{\partial v_0} = -\frac{\gamma l(l-1)^{\frac{1}{2}}}{l^{\frac{1}{2}}} - \frac{\delta(l-1)^{\frac{1}{2}}}{l^{\frac{1}{2}}}\frac{(1-\gamma v_0/\lambda^3)}{(1+\delta v_0/\lambda^3)}C_{(l-1)}. \tag{16}$$

Equating terms of order unity from Eqs. (15) and (16) we get

$$\gamma = 1+1.61\frac{(1-\gamma v_0/\lambda^3)}{(1+\delta v_0/\lambda^3)}, \tag{17}$$

while order $1/l$ gives,

$$\tfrac{3}{2}\gamma - \frac{(1-\gamma v_0/\lambda^3)}{(1+\delta v_0/\lambda^3)} = 6.12\frac{(1-\gamma v_0/\lambda^3)}{(1+\delta v_0/\lambda^3)}. \tag{18}$$

Equation (17) is satisfied by putting $\gamma=-\delta=2.61$. If we put $\gamma=-\delta$ in Eq. (18) we find $\gamma=2.44$. Thus we may say that to within 7 percent,

$$C_l = (1/l^{\frac{5}{2}})\{\lambda^3(1-2.61v_0/\lambda^3)\}^{l-1} \tag{19}$$

is a solution of Eq. (12) for large l.

We see that v_0 in Eq. (17) is multiplied by 2.61. This number arises as a result of the summation $\sum 1/l^{\frac{5}{2}}$, which receives contributions from a large number of terms. What has happened is that the λ linking between two adjacent particles is not only affected by the Q linking directly between them but by those through other particles as well, since the particles are correlated in position by the Gaussian factors. It is true that this correlation drops off fairly rapidly as we move away from a given particle to its neighbor, next nearest neighbor, etc., but only in the ratio as the terms of the series $1/l^{\frac{5}{2}}$.

In the above calculation C_l has been evaluated in the limit of infinite volume (and zero "packing"). Thus the only effect which has been taken account of has been that of the repulsive core in opposing the attractive force of the statistics. The modification to (19) as a result of the packing effect is considered in subsection (c).

(c) Statistical Method

We would now like to verify, at least to first order, the previously obtained result for S_P, and furthermore show that the term $\exp(\sum l^2 m_l v_0/2V)$ appearing in it, should be cancelled if the complete packing effect is considered. That is to say that C_l would have been multiplied by $\exp(-l^2 v_0/2V)$ if we had not computed it in the limit of zero packing.

We thus return to the consideration of S_P, but now divide Q into two products, one involving interactions between those particles of the same cluster that are *significantly correlated by the statistics*, as discussed above, and the other involving all other interactions. We furthermore make a transformation of coordinates as follows. In each cluster let some particle be number 1, then $\mathbf{r}_1=\mathbf{R}$. The other $(l-1)$ coordinates are chosen by letting $\mathbf{r}_1-\mathbf{r}_2=\mathbf{y}_1$, $\cdots \mathbf{r}_{(l-1)}-\mathbf{r}_l=\mathbf{y}_{(l-1)}$, etc. It is convenient to introduce one additional coordinate for each cluster $\mathbf{y}_l=\mathbf{r}_l-\mathbf{r}_1$. To do this we must introduce the Dirac delta function,

$$\frac{1}{(2\pi)^3}\int d\mathbf{p}d\mathbf{y}_l \exp\{i\mathbf{p}\cdot(\mathbf{y}_1+\mathbf{y}_2+\cdots+\mathbf{y}_{l-1}-\mathbf{y}_l)\},$$

if we are not to overspecify the transformation. Thus,

$$S_P = \int \prod_{\phi}^{\text{clusters}} d\mathbf{p}_\phi d\mathbf{R}_\phi d\mathbf{y}_1\cdots d\mathbf{y}_N \times\exp\{i\mathbf{p}_\phi\cdot(\mathbf{y}_1{}^\phi+\cdots+\mathbf{y}_l{}^\phi)\} \times\exp\left\{-\frac{\pi}{\lambda^2}[(y_1{}^\phi)^2+\cdots]\right\} \times\prod_{i=1}^{N}\prod_{\nu}^{\text{correlated}} q(|\mathbf{y}_i+\mathbf{X}_{\nu i}|) \times\prod_{\mu}^{\text{uncorrelated}}[1-v_0\Delta(|\mathbf{y}_i+\mathbf{X}_{\mu i}|)], \tag{20}$$

where we have used the index ϕ to designate the clusters. The vector $\mathbf{X}_{\mu i}$ is the sum of the y's and R's necessary to locate the μth particle, so that $\mathbf{y}_i+\mathbf{X}_{\mu i}$ is the relative position vector between the μth and ith particles We note that the index νi designates those particles correlated to the ith particle by the statistics while μi refers to all the others. Furthermore, since the index i runs from one to N, μ may take on at most $(N-1)/2$ values so as not to repeat an interaction twice.

On carrying out the product $\prod_\mu$ to first order in v_0 we find,

$$\prod_\mu[1-v_0\Delta(|\mathbf{y}_i+\mathbf{X}_{\mu i}|)]\rightarrow 1-v_0\sum_\mu\Delta(|\mathbf{y}_i+\mathbf{X}_{\mu i}|). \quad (21)$$

We now argue that for almost all values of the set of coordinates $\{\mathbf{y}_1,\cdots\mathbf{y}_N,\mathbf{R}_\phi\}$, the coordinates $\mathbf{X}_{\mu i}$ will be randomly distributed since there is no correlation between the ith and μth particles. Hence we may take

$$\sum_\mu f(\mathbf{X}_{\mu i}) \quad \text{into} \quad \int dn_\mu f(\mathbf{X}_{\mu i})=\frac{\rho}{2}\int d\mathbf{X}_{\mu i} f(\mathbf{X}_{\mu i}).$$

We take ρ since the number of particles correlated to i is finite and negligible compared to N. The factor 2 arises because of the previously mentioned restriction on μ. Thus $v_0\sum_\mu\Delta(|\mathbf{y}_i+\mathbf{X}_{\mu i}|)\rightarrow v_0\rho/2$.

At this point we note that we need not have used the Dirac delta function approximation for this part of the problem, but could have let $q(r_{ij})=1-g(r_{ij})$, where g is any function whose volume integral is v_0. However, we needed this approximation for evaluating the C_l's. We also see that for this statistical approximation to be valid we require that $\rho v_0/2\ll 1$, which is consistent with the requirement on the delta function approximation.

The term $(1-\rho v_0/2)$ may now be taken out from under the product $\prod_{i=1}^{N}$ to give

$$S_P=(1-\tfrac{1}{2}\rho v_0)^N\int\prod_\phi d\mathbf{p}_\phi d\mathbf{R}_\phi d\mathbf{y}_1\cdots d\mathbf{y}_N$$

$$\times\exp\{i\mathbf{p}_\phi\cdot(\mathbf{y}_1{}^\phi+\cdots)\}\exp[-(\pi/\lambda^2)\{(y_1{}^\phi)^2+\cdots\}]$$

$$\times\prod_i\prod_\nu q(|\mathbf{y}_i+\mathbf{X}_{\nu i}|). \quad (22)$$

The integral of (22) is the same as A_P, except that now any one particle is linked by Q only to the *finite* number of particles with which it is significantly correlated by the statistics. The remainder of the internal linkings, which occur for very large clusters, have been taken into account and, as is to be expected, have resulted in the elimination of the multiplying factor $\exp\{\sum_l v_0 l^2 m_l/2v\}$ of Eq. (8). The factor $(1-\frac{1}{2}\rho v_0)^N$ of (22) is of course the same as the factor $\exp(-\frac{1}{2}\rho v_0 N)$ of (8) to first order (since we shall eventually be interested in its logarithm).

The main effects of the finite volume V of the container have now been taken care of, and have produced the factor $(1-\frac{1}{2}\rho v_0)^N$ of (22) [or $\exp(-\frac{1}{2}\rho v_0{}^N)$ of (8)]. In general the limits of the remaining integral of (22) may now be made infinite and its evaluation then proceeds exactly as carried out in II (b). To facilitate the discussion of very low temperatures (below the Bose-Einstein condensation point) however, it is convenient to evaluate the volume dependence of this integral to order $1/V$. This applies also to the case of the perfect gas, and has been carried out in the Appendix. It is there found that C_l becomes proportional to

$$\{1/l^3+(\lambda^3-2.61v_0)/V\}$$

rather than merely $1/l^3$ as derived in (b) [Eq. (19)].

III. EVALUATION OF THE PARTITION FUNCTION AND THERMODYNAMIC PROPERTIES

The partition function has now been reduced to the form

$$Z=[1/(N!\lambda^{3N})]\sum_P S_P, \quad (23)$$

where S_P is characterized by the number of clusters m_l of size l:

$$S_P=\exp\{-\tfrac{1}{2}\rho v_0 N\}\prod_l (VC_l)^{m_l}. \quad (24)$$

It is convenient to define a quantity b_l given by

$$lb_l=C_l=\{1/l^3+(\lambda^3-2.61v_0)/V\}\times\{\lambda^3(1-2.61v_0/\lambda^3)\}^{l-1}. \quad (25)$$

One now sees[2] that Eq. (23) becomes:

$$Z=\frac{\exp(-\frac{1}{2}\rho v_0 N)}{\lambda^{3N}}\sum_{\{m_l\}}{}'\prod_{l=1}^{N}\frac{[b_l V]^{m_l}}{m_l!}. \quad (26)$$

The sum in Eq. (26) is over all sets of numbers $\{m_l\}$ such that $\sum_l lm_l=N$. This sum may now be evaluated by the method of reference 2.[3] However, the results will differ from those of this latter paper in that we will always obtain an equation for the density N/V, and furthermore this equation will always have solutions. This of course is due to the fact that we have included the volume dependence of C_l. In the limit as v_0 goes to zero (ideal gas), this treatment is identical to that of London.[4]

Having obtained the partition function, one may determine the thermodynamic properties in the usual way. They are found to be continuous functions of the temperature T. However, an examination of the specific heat C_v shows that there exists a point T_e for which it exhibits a discontinuity in its slope. This point is identified as a transition point. It occurs at a lower temperature than the corresponding transition for the ideal gas.

[3] Since we have derived a general form for C_l [Eq. (25)] which is valid for large l only, corrections must also be made to the C_l's for small l. Such corrections can of course quite readily be evaluated but turn out to be surprisingly small and may in general be ignored.

[4] F. London, Phys. Rev. **54**, 947 (1938).

One also finds that here the pressure is dependent upon the volume below as well as above the transition point. This occurs because of the factor $\exp(-\frac{1}{2}\rho v_0 N)$ which appears as a coefficient in the expression for Z.

Both the partition function and the specific heat take on simple forms [Eqs. (27) and (28) below] below the transition point. Above the transition point these functions are expressed in the usual type of series form, obtained by solving the density equation for the fugacity, and inserting this result into $\ln Z$.

Thus for $T<T_c$,

$$\ln Z=-\frac{\rho v_0 N}{2}+\frac{1.31V}{\lambda^3-\lambda_0^3}+N\ln\left[\frac{\lambda^3-\lambda_0^3}{\lambda^3}\right], \quad (27)$$

and

$$C_v=\frac{(C_v)_I}{x^2}\left\{\frac{6}{5}\left(\frac{1}{x}-\frac{1}{x_c}\right)+1-\frac{4}{5}\frac{x}{x_c}(1-x_c)\right\}, \quad (28)$$

where $(C_v)_I$ is the specific heat of the ideal Bose-Einstein gas, $\lambda_0^3=2.61v_0$, and $x=(1-\lambda_0^3/\lambda^3)$; x_c is the same quantity evaluated at the condensation temperature. The simple factors 6/5 and 4/5 arise because we have set

$$\frac{\sum_l 1/l^{5/2}}{\sum_l 1/l^{5/2}}=\frac{2.61}{1.31}\simeq 2. \quad (29)$$

We would again like to caution the reader that the above results are not valid for temperatures so low that the long-range density fluctuations become important,[1] nor for densities so high that $\rho v_0 \geq \frac{1}{2}$.

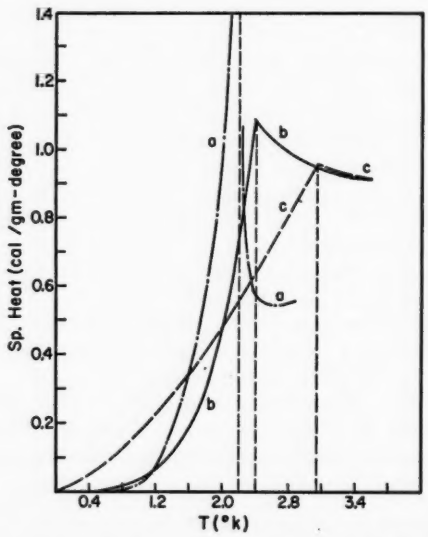


FIG. 1. The specific heat of He^4 at constant volume. Curve (*a*) shows the experimental results for liquid helium, (*b*) is obtained from the present theory, while (*c*) is obtained from the ideal gas theory.

IV. COMPARISON WITH LIQUID HELIUM AND THE IDEAL GAS

The above results have been applied using a value of r_0 of 1.8 A. Inserting the density of liquid helium at the λ-point, we obtain a curve for the specific heat at constant volume as shown in Fig. 1. For purposes of comparison we have also included the curve for the ideal gas and that experimentally obtained for the specific heat of liquid helium under its saturated vapor pressure. It is seen that the transition temperature obtained here is very much closer to that observed than is predicted by the treatment of the perfect Bose-Einstein gas. The rate of decrease immediately below the transition point approximately follows a T^6 law. For temperatures near zero, however, it has become a T^3 law.

APPENDIX

We now wish to show that the $1/l^{\frac{3}{2}}$ that occurs as a coefficient in Eq. (19) for C_l should be replaced by $\{1/l^{\frac{3}{2}}+(\lambda^3-2.61v_0)/V\}$. We shall first consider the case of the ideal gas (i.e., $v_0=0$) where the same situation arises.

The integral to be evaluated is

$$VC_l{}^P=VC_l(v_0=0)=\int d\mathbf{r}_1\cdots d\mathbf{r}_l \times\exp\{-(\pi/\lambda^2)[r_{12}{}^2+\cdots+r_{l1}{}^2]\}. \quad (1A)$$

For small values of l one may extend the limits of integration to infinity which gives,

$$C_l{}^P=\frac{1}{l^{\frac{3}{2}}}\lambda^{3(l-1)}. \quad (2A)$$

We shall now attempt to evaluate the integral of Eq. (1A) to lowest order in λ^3/V for the case $\lambda^3/V\gg 1/l^{\frac{3}{2}}$. The limits of integration are now such that each coordinate is restricted to within the volume V. Let us first make the transformation $\mathbf{y}_1=\mathbf{r}_1$, $\mathbf{y}_2=\mathbf{r}_2-\mathbf{r}_1$, $\cdots\mathbf{y}_l=\mathbf{r}_l-\mathbf{r}_{(l-1)}$.

The new limits of integration will actually be functions of the coordinates themselves, but since we are only interested in the lowest order in λ^3/V, we may take them so that each of the $\mathbf{y}_j$'s is independently restricted to within the volume V. This permits us to immediately integrate over $\mathbf{y}_1$.

$$VC_l{}^P=V\int d\mathbf{y}_2\cdots d\mathbf{y}_l\exp\{-(\pi/\lambda^2)[y_2{}^2+y_3{}^2\cdots +y_l{}^2+|\mathbf{y}_2\cdots+\mathbf{y}_l|^2]\}. \quad (3A)$$

It is now convenient to replace the position vector $\mathbf{y}_2\cdots+\mathbf{y}_l$ occurring in the bracket by a new coordinate $\mathbf{x}$. To do this we must introduce the Dirac delta function $\delta(\mathbf{y}_2\cdots+\mathbf{y}_l+\mathbf{x})$, integrating $\mathbf{x}$ over the volume V. However, we must be careful to correctly represent the delta function. The closure theorem here requires

the use of plane waves normalized in a box of volume V, since these obey boundary conditions appropriate to the limits of integration. Thus we set $\delta(\mathbf{y}_2+\cdots\mathbf{x}) = (1/V)\sum_{\mathbf{k}} \exp\{i\mathbf{k}\cdot(\mathbf{y}_2+\cdots\mathbf{x})\}$. Substituting these results into Eq. (3A), we find

$$C_l^P=\frac{1}{V}\sum_{\mathbf{k}}\left\{\int d\mathbf{y}\,\exp[i\mathbf{k}\cdot\mathbf{y}-(\pi/\lambda^2)y^2]\right\}^l. \quad (4A)$$

We may now take $\sum_{\mathbf{k}} f(\mathbf{k})$ into $[V/(2\pi)^3]\int f(\mathbf{k})d\mathbf{k} + f(0)$. The term $f(0)$ must be added because the integral form gives the state $k=0$ zero weight, thereby throwing away its contributions. Thus,

$$C_l^P=\frac{1}{V}\left\{\int d\mathbf{y}\,\exp[-(\pi/\lambda^2)y^2]\right\}^l + \frac{1}{(2\pi)^3}\int d\mathbf{k}\left\{\int d\mathbf{y}\,\exp(i\mathbf{k}\cdot\mathbf{y}-(\pi/\lambda^2)y^2)\right\}^l. \quad (5A)$$

Upon evaluation, the first term becomes $(\lambda^{3l}/V)\times\{P(\pi^{\frac{1}{2}}V^{\frac{1}{3}}/2\lambda)\}^{3l}$, where $P(x)$ is the probability integral so normalized that $P(\infty)=1$.

It is quite straightforward to show that the second integral can be rewritten as

$$\frac{\lambda^{3l}}{(2\pi)^3}\left\{\int_{-\infty}^{\infty} dk[\exp(-lk^2\lambda^2/4\pi)] \times\left(P(\pi^{\frac{1}{2}}V^{\frac{1}{3}}/2\lambda)+\frac{2}{\pi^{\frac{1}{2}}}\exp(-\pi V^{\frac{1}{3}}/4\lambda^2) \times\int_0^{k\lambda/2\pi} dx\,\exp(x^2)\sin(\pi^{\frac{1}{2}}V^{\frac{1}{3}}x/\lambda)\right)^l\right\}^3.$$

The second term in the brackets may be dropped. The integral itself has a very small value because of the highly oscillatory nature of the integrand, in addition to which it is multiplied by an exponential which is extremely small in value. Hence this term becomes $(\lambda^{3(l-1)}/l^{\frac{3}{2}})\{P(\pi^{\frac{1}{2}}V^{\frac{1}{3}}/2\lambda)\}^{3l}$. Thus,

$$C_l^P=[\lambda P(\pi^{\frac{1}{2}}V^{\frac{1}{3}}/2\lambda)]^{3(l-1)}P^3(\pi^{\frac{1}{2}}V^{\frac{1}{3}}/2\lambda) \times\{\lambda^3/V+1/l^{\frac{3}{2}}\}. \quad (6A)$$

Now since $V^{\frac{1}{3}}/\lambda$ is very large, we may set $P=1$. Its only effect otherwise is to multiply λ^3 by a number very slightly less than unity. Hence, we may take

$$C_l^P=\lambda^{3(l-1)}\{\lambda^3/V+1/l^{\frac{3}{2}}\}. \quad (7A)$$

This form of C_l^P is to be used for large l as well as small since in the latter case $\lambda^3/V \ll 1/l^{\frac{3}{2}}$. The use of Eq. (7A) is identical to the procedure employed by London.[3]

We shall now return to a consideration of Eq. (10), using the results obtained for the perfect gas as a guide. Let us again introduce new coordinates along with the Dirac delta function (appropriately expressed) to obtain

$$C_l=\frac{1}{V}\int d\mathbf{y}_1\cdots d\mathbf{y}_l\,\exp\{-(\pi/\lambda^2)[y_1^2+\cdots+y_l^2]\} \times Q_l'(\mathbf{y}_1,\cdots\mathbf{y}_l)+\frac{1}{(2\pi)^3}\int d\mathbf{k}\int d\mathbf{y}_1\cdots d\mathbf{y}_l \times\exp\{-(\pi/\lambda^2)[y_1^2+\cdots y_l^2] +i\mathbf{k}\cdot(\mathbf{y}_1+\cdots\mathbf{y}_l)\}Q'. \quad (8A)$$

Here $Q'(\mathbf{y}_1,\cdots\mathbf{y}_l)$ is the same as $Q(\mathbf{r}_1,\cdots\mathbf{r}_l)$ but expressed in terms of the y's. We now recall that in the case of the ideal gas the fact that the y's were restricted to within a volume V made no difference in the final result (to the order in which we were interested) as far as the second term of (5A) was concerned. There is no reason to believe that the same should not be true of the analogous second term of (8A) once the contribution to the "packing effect" has been considered. We therefore extend the limits of integration to ∞. The integration over $\mathbf{k}$ then becomes a delta function and this term gives $(1/l^{\frac{3}{2}})[\lambda^3(1-2.61v_0/\lambda^3)]^{l-1}$ as previously obtained.

In treating the first term, we make use of the observation that $Q(\mathbf{r}_1,\cdots\mathbf{r}_l)$ behaves very much as if it operated only between neighboring particles in the cycle formed by the λ linkings, but with an effective excluded volume of $2.61v_0$. This term then becomes $(1/V)\{[P(\pi^{\frac{1}{2}}V^{\frac{1}{3}}/2\lambda)\lambda]^3-2.61v_0\}^l$ which on taking $P\to 1$ gives $(1/V)\{\lambda^3(1-2.61v_0/\lambda^3)\}^l$. Thus finally,

$$C_l=\left\{\frac{1}{l^{\frac{3}{2}}}+\frac{(\lambda^3-2.61v_0)}{V}\right\}\left\{\lambda^3\left(1-\frac{2.61v_0}{\lambda^3}\right)\right\}^{l-1}. \quad (9A)$$

New Method for Measuring the α-Particle Range and Straggling in Liquids

GUNNAR ANIANSSON
Division of Physical Chemistry, The Royal Institute of Technology, Stockholm, Sweden
(Received October 25, 1954)

The ranges of Po α particles in water and five other liquids have been measured with a new method involving a steel surface as support for the α source and a $CaWO_4$ crystal surface as detector for the α particles. The surfaces are optically plane and parallel and constitute the walls of a microcell for the liquid. The range obtained at 25°C in water was 39.9 microns, 2.7 percent *longer* than the value 38.8 microns, computed with the Bragg law from range measurements in gaseous hydrogen and oxygen. In the other five liquids the ranges obtained also were longer (2.3–4.6 percent) than the correspondingly computed Bragg law values.

THE range of RaC′ α particles (energy: 7.68 Mev) in water was reported by Philipp[1] in 1923 to be 13 percent shorter than what could be computed from the measured range in water vapor. This was supposed to be caused by the strong association of the water molecules in the liquid state. Nonassociative liquids such as benzene and pyridine showed similar effects of only 5 percent when measured in the same apparatus. Michl[2] in 1914 reported measurements with a photographic method on Po α-particle (energy: 5.30 Mev) ranges in several liquids. The range obtained for water was 32 microns, which is 17 percent less than the value 38.7 microns given by the Bragg law when one uses for hydrogen the stopping power value 0.224 and for oxygen 1.060 relative to air[3] and a mean range in air of 38.42 mm at 15°C and 1 atmos.[4] More recent measurements by Appleyard[5] and de Carvalho[6] on Po α particles gave 33.2 and 38.1 microns respectively as the range in water. Appleyard's value is 14 percent and de Carvalho's value 1.5 percent smaller than that predicted by Bragg's law. A redetermination of the α-particle range in water and other liquids would therefore seem to be of interest.

The new method[7] is built on the following principle. The liquid is placed between two plane, parallel surfaces constituting the walls of a microcell. One wall is the support of an α-particle source located at its center and of negligibly small thickness while the other is a scintillation crystal. Those α particles which are sent out within a certain angle to the normal reach the scintillation crystal and give light flashes which are counted by a photomultiplier.

The counting rate *versus* the distance between the surfaces is given by the following formula:

$$N=\int_x^\infty \frac{N_0}{2\alpha\sqrt{\pi}}\exp[-(R-R_0)^2/\alpha^2]\left(1-\frac{x}{R}\right)dR, \quad (1)$$

where N_0 is the total number of disintegrations per unit time in the source, x the distance between the surfaces, R_0 the mean range, α the straggling parameter, and R the variable in the (Gaussian) straggling function $(1/2\alpha\sqrt{\pi})\exp[-(R-R_0)^2/\alpha^2]$ assumed to be valid.

In the limit as $\alpha/R_0\to 0$, i.e., for a vanishingly small straggling, Eq. (1) reduces to

$$N=\tfrac{1}{2}N_0(1-x/R_0) \text{ for } x\leq R_0 \text{ and } N=0 \text{ for } x\geq R_0, \quad (2)$$

i.e., N as a function of x tends to a straight line intersecting the x-axis at R_0. For values of α of one or a few percent of R_0, the same will be very nearly true for (1) except for x-values within $R_0-\alpha$ and $R_0+\alpha$, where the sharp bend at $x=R_0$ is replaced by a gradual change in the slope of the curve. The value of N at $x=R_0$ is to a good approximation given by $(N_0\alpha/4R_0\sqrt{\pi})$, from which α may be obtained.[8]

The experimental arrangement is shown schematically in Fig. 1. A piston of 7-mm diameter was ground to a very good fit into a cylinder body, both being made from a type of steel used for gauge blocks. The end surfaces of the cylinder and the piston were ground plane[9] to within about one-tenth of a micron and controlled in a

[1] K. Philipp, Z. Physik **17**, 23 (1923).
[2] W. Michl, Sitzber. Akad. Wiss. Wien, Math. naturw. Kl. **123**, 1965 (1914).
[3] L. H. Gray, Proc. Cambridge Phil. Soc. **40**, 87 (1944).
[4] M. G. Holloway and M. S. Livingston, Phys. Rev. **54**, 18 (1938).
[5] R. K. Appleyard, Nature **163**, 527 (1949); Proc. Cambridge Phil. Soc. **47**, 443 (1951).
[6] H. G. de Carvalho, Phys. Rev. **78**, 330 (1950); H. G. de Carvalho and H. Yagoda, Phys. Rev. **88**, 273 (1952).
[7] G. Aniansson, Report at the Instruments and Measurements Conference, Stockholm, 1952 [Arkiv Fysik **7**, 74 (1954)].

[8] It may be remarked that instead of (1), one may write in general:

$$N=\int_x^\infty \tfrac{1}{2}N_0S(R)\left(1-\frac{x}{R}\right)dR, \quad (3)$$

where $S(R)$ is any straggling function. With given $N(x)$, one obtains $S(R)$ from

$$S(R)=(2x/N_0)(d^2N/dx^2). \quad (4)$$

With the aid of a differential discriminator, one may also select pulses under a certain amplitude for counting, resulting in a relation between the count rate n and the distance x of the form

$$n=\int_x^\infty CS(R)\frac{x}{R^2}dR; \quad C=\text{a constant} \quad (5)$$

and

$$S(R)=\frac{1}{C}\left(n-x\frac{dn}{dx}\right). \quad (6)$$

Equation (6) has the advantage over (4) that the derivative is of only first order.

[9] C. E. Johansson, AB, Eskilstuna, Sweden.

Michelson interferometer. The parallelepipedal (3×3 =15 mm) $CaWO_4$ scintillation crystal was ground plane on one side to the same tolerance.[10] The crystal is slid on to the cylinder end surface to close contact which could be controlled by the light interference fringes that appear. The upper surface of the crystal is coarse-ground to diminish the trapping of light through total reflection. A glass-covered conical light collector brings the light to the photocathode of an EMI 6260 multiplier tube and serves as a container for the liquid which has free access to the space between the crystal and the piston (the piston diameter is 7 mm, the crystal breadth 3 mm).

The piston is moved by a lever connected to the cylinder and the piston by two plane metal springs. The displacement is read on a "mikrokator"[11] calibrated to within 0.05 micron.[12] The distance between the piston and cylinder end surfaces at the zero point of the "mikrokator" was determined with a second "mikrokator." The crystal, light collector and photomultiplier were removed during this operation. The polonium α source was evaporated from a platinum support on to the center of the piston end surface. An autoradiograph showed that the polonium was evenly deposited on a circular area of about 1 mm diameter in the center of the piston end surface.

The pulses from the multiplier were amplified 30 times, passed through a discriminator which excluded pulses under a certain amplitude—among them thermionic pulses from the phototube—and counted with a scaler. The higher the discriminator level is, the longer must the α particle travel in the scintillation crystal to produce the necessary light amount. This results in a correspondingly smaller apparent range, and thus range measurements were made with different discriminator levels and the ranges obtained were extrapolated to zero discriminator level.

Figure 2 shows the results obtained with absolute ethyl alcohol with discirminator settings of 5, 10, 15, 20, and 40 volts. The curves are straight lines except in a small region near the apparent range as predicted by (1). The plot of apparent range *versus* discriminator level is also a straight line, which facilitates the extrapolation to zero level.

With the same crystal, light collector, and counting equipment, the apparent mean range in dry air was determined for different discriminator levels and extrapolated to zero level. The result, obtained at 25.0°C and 748.4 mm Hg, was 39.80 mm which at the normal temperature 15°C and pressure 760 mm Hg would correspond to 37.87 mm. This is 0.55 mm or 1.4 percent

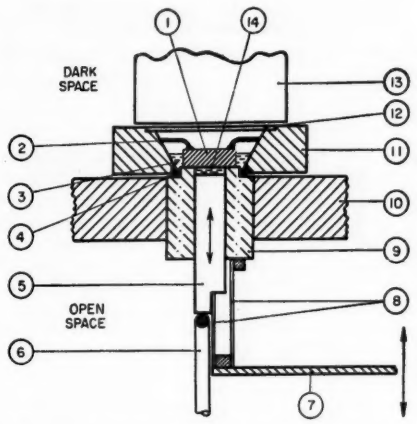


FIG. 1. The experimental arrangement. (1) Scintillation crystal, $CaWO_4$, 3×3×15 mm. (2) Metal spring pressing the crystal against the cylinder end surface. (3) Liquid. (4) Gasket. (5) Piston. (6) "Mikrokator" measuring arm. (7) Lever. (8) Plane metal springs. (9) Cylinder. (10) Base plate. (11) Light collector and liquid container. (12) Glass plate. (13) EMI 6260 photomultiplier tube. (14) Po α source, diameter 1 mm.

shorter than the accepted standard value 38.42 mm obtained with ionization methods and is caused partly by the thin aluminum layer which for reasons given below was evaporated onto the middle part of the ground surface of the crystal. The rest of this discrepancy may be caused by a thin insensitive surface layer which has probably melted during the grinding process.

According to present theories[13,14] of the scintillation process in solutions in liquid, the energetic particle gives off its energy to the solvent from which by some mechanism it is transferred to the solute molecules. If this transfer mechanism were efficient over distances of some tenths of a micron for the particular system,

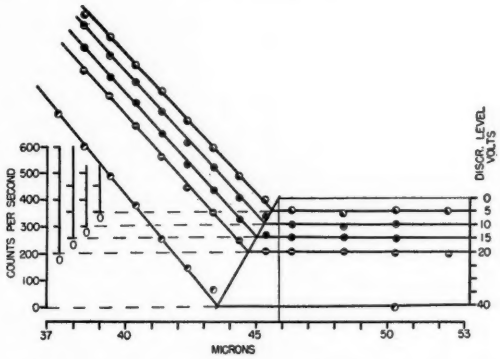


FIG. 2. Counting rate *versus* distance curves for absolute ethyl alcohol. The parallel, vertical displacements of the curves are proportional to the discriminator level.

[10] Junger, AB, Stockholm, Sweden.
[11] Type 509–7, manufactured by C. E. Johansson, AB.
[12] When still higher accuracy is sought, interference or other optical methods may be used. The distance between the piston end and the crystal surface might for example be directly measured through the interference of monochromatic light rays reflected at the two surfaces.

[13] S. C. Curran, *Luminescence and the Scintillation Counter* (Butterworths Scientific Publications, London, 1953).
[14] J. B. Birks, *Scintillation Counters* (Pergamon Press, London, 1953).

TABLE I. Experimental values, corrections, and comparison with the Bragg law.

Liquid	Apparent mean range extrapol. to zero discr. level microns	Corrected mean range (average) microns	Density g cm^{-3}	Stopping power exptl	Stopping power theor[a]	$\frac{s_{theor}}{s_{exptl}} = \frac{R_{exptl}}{R_{theor}}$
Water	39.25 39.25 39.17	39.88	0.9969	1.469	1.508	1.027
Ethyl alcohol	45.85 45.90	46.53	0.7866	4.07	4.198	1.030
Benzene	43.40 43.55	44.09	0.8727	6.56	6.714	1.023
Pyridine	40.40	40.97	0.9775	6.39	6.585	1.031
n-heptane	49.90 49.65 49.75	50.48 50.45	0.6790 0.6792[b]	9.46	9.849	1.042
Iso-octane (2,2,4-trimethyl pentane)	49.40 49.75	50.27	0.6871	10.70	11.19	1.046

[a] $s_O = 1.060$, $s_H = 0.224$, and $s_C = 0.895$ according to reference 3.
[b] From a different source.

liquid—$CaWO_4$, it would lead to a noticeable increase in the apparent range. Measurements were therefore made on two liquids, water and benzene, (a) with the crystal uncovered and (b) with the crystal covered by a thin but opaque aluminum film which would greatly reduce any energy transfer from the liquid to the crystal. The ranges found under (a) were 0.3 to 0.5 micron longer than those found under (b). Allowing for the contributing effect of the aluminum film, the differences become still smaller and completely within the limits of experimental error. For safety the film was, however, allowed to remain in the subsequent measurements and all the results reported have been obtained with this film on the crystal.

There is thus only one correction to apply to the "zero-level" ranges obtained, i.e., the one caused by the thin aluminum film and the insensitive layer on the crystal surface. In air the reduction in range caused by these layers is 1.4 percent and the same reduction applies for the ranges in the liquids provided the alteration with α-particle energy of the stopping power of the liquids relative to air may be neglected. There is no reason to take such an alteration into account in this case since the correction is almost as small as the experimental error.

In Table I are given the results obtained for six liquids at 25°C. The densities were determined with a pycnometer at the same temperature. The stopping power was computed from the formula

$$s = (38.42 \times 1.226 \times 10^{-3}/14.56) \times (M/\rho R_0), \qquad (7)$$

where 38.42 mm is the standard range in air of density 1.226×10^{-3} g cm^{-3} and mean atomic weight 14.56, M is the molecular weight of the liquid, ρ its density in g cm^{-3}, and R_0 is the corrected range in the liquid in mm.

It is seen that the largest difference between any two experimental values for the range for the same liquid is less than 1 percent which is the assumed value for the uncertainty in the measurements.

The ranges are all *longer* than those computed from the Bragg law and measurements on gases, thus giving the impression of a general density effect.[15] In contrast to this the aforementioned results by Michl, Philipp, Appleyard, and de Carvalho were all *lower* than the values predicted by the Bragg law. The differences between their results for water and that in the present work are 20 percent, 16 percent, 17 percent, and 4 percent respectively, clearly outside the assumed limit of accuracy of the present method.

Of the six compounds listed in Table I, only water has been subject to range measurements in the gaseous state with Po α particles. For water vapor, Förster[16] obtained a 3 percent longer range than that predicted by the Bragg law, in agreement with the present result, whereas de Carvalho and Yagoda[6] in a single measurement obtained a range 2.5 percent lower than that predicted by the Bragg law, in approximate agreement with Philipp's[1] result for RaC′ α particles.

ACKNOWLEDGMENTS

The author wishes to express his thanks to the head of the division, Professor Ole Lamm, for suggesting and supporting this investigation. Thanks are also due to the members of the division for valuable discussions and other aid. The investigation was financed by the Swedish Atomic Committee.

[15] R. M. Sternheimer, Phys. Rev. **93**, 351 (1954).
[16] M. Förster, Ann. Physik **27**, 372 (1936).

Mechanism of the Termination of the Geiger Plateau Region*

H. L. WISER AND A. D. KRUMBEIN
Physics Department, University of Maryland, College Park, Maryland
(Received November 8, 1954)

The mechanisms responsible for the termination of the Geiger counter plateau region at its high-voltage end were investigated by a direct experimental method. Geiger counters were operated in a normal manner and counting rate *versus* voltage and number of positive ions per pulse *versus* voltage were measured in the breakdown region. Analysis of the data indicated that spurious counts could result from two phenomena: (1) Secondary electrons created by positive-ion bombardment of the cathode; and (2) metastable atom and molecule decay in the gas. The rate of spurious counts caused by positive-ion bombardment was found to be proportional to the rate of positive ions striking the cathode and, therefore, proportional to the voltage. This accounted for the plateau slope, if present, and the slight increase in slope at the beginning of the breakdown region. It is shown that the exponential rise in counting rate in the breakdown region can be explained by the decay of metastable atoms and molecules that exist after the deadtime of the primary pulses. Experimental results indicate that metastable atoms were dominantly responsible for spurious counts at low pressures (about 40 mm Hg) and metastable molecules at high pressures (200 mm Hg).

1. INTRODUCTION

SPURIOUS pulses in Geiger counters conceivably may be caused by positive ions, negative ions, or photons. The theory of positive-ion bombardment of the cathode as a source of spurious counts has been presented by Korff and Present.[1] However, their theory could not explain the relatively sudden onset of spurious counts with the increase of applied voltage. Willard and Montgomery's explanation that spurious counts were caused by negative ions could not be accepted for the breakdown region where the counting rate was rapid.[2] Colli and Facchini reported photon bombardment of the cathode to be responsible for the continuous discharges seen in their counters.[3] Their counters, however, had no plateau region and were operated at voltages that would have been well above the start of the breakdown region had there been plateaus. DuToit conducted two different experiments in an attempt to settle the controversy, but his results were indecisive, one experiment favoring the positive-ion theory and the second favoring the photon theory.[4] In the present study a direct experimental approach was used for the purpose of allowing the experiment itself to provide data to resolve this dilemma. Geiger counters were operated in a normal manner and counting rate *versus* voltage and number of positive ions per pulse *versus* voltage were determined in the breakdown region at small increments of applied voltage.

II. APPARATUS AND EXPERIMENTAL PROCEDURE

Demountable counters were used throughout. The whole counter, shown in Fig. 1, was encased in an evacuated glass chamber and filled with appropriate gases at the desired pressures. The cathodes, shown in Fig. 2, were designed so as to maintain the same surfaces for various radii, thereby eliminating surface structure as a variable parameter. This had been a parameter of unknown behavior responsible for discrepancies in the results of different experimenters.[5] A width of brass sheet (0.002 in. thick) was cut to equal the circumference of the largest of the desired cathodes, so that, when rolled into a cylinder, the springy brass held itself firmly against the inside of a glass tube of equal circumference, forming a closed metal surface.

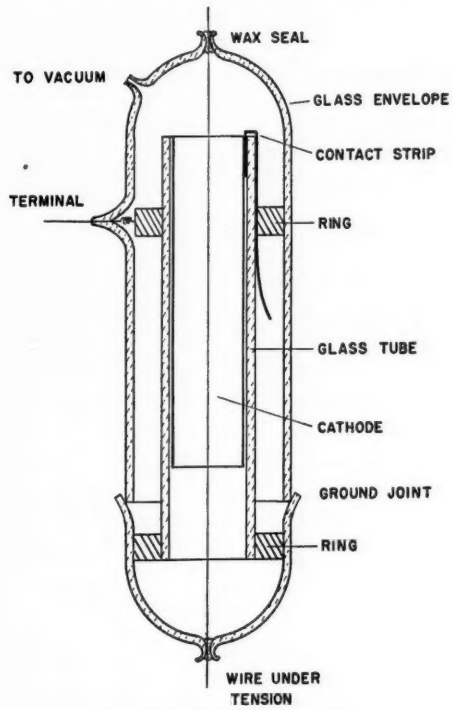


FIG. 1. Geiger counter assembly.

* Supported by U. S. Bureau of Ships. The information contained in this paper is part of a thesis submitted by one of the authors, H. L. Wiser, in partial fulfillment of the requirements for the Ph.D. degree at the University of Maryland.

[1] S. A. Korff and R. D. Present, Phys. Rev. **65**, 274 (1944).
[2] D. Willard and C. G. Montgomery, Rev. Sci. Instr. **21**, 520 (1950).
[3] L. Colli and U. Facchini, Phys. Rev. **88**, 987 (1952).
[4] S. J. DuToit, Phys. Rev. **73**, 1473 (1948).
[5] L. B. Loeb, *Fundamental Processes of Electrical Discharges in Gases* (John Wiley and Sons, Inc., New York, 1939), p. 378.

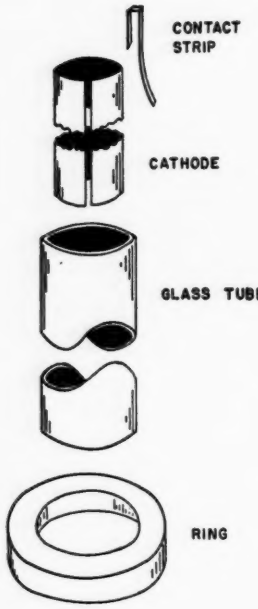


FIG. 2. Cathode assembly.

When the counters were disassembled for experiments with smaller cathode radii, the same piece of brass sheet was cut to fit the inner diameter of the next smaller glass tube. This process was continued for subsequent runs with still smaller cathode radii.

The tungsten anode wires, which were 3, 4, or 8 mils in diameter, were changed after only two or three runs because their surfaces were damaged by the intense electron bombardment, giving rise to erratic counter behavior and nonreproducible results in the breakdown region.

Each experimental run was preceded by the cleaning, evacuating, and filling of the Geiger counters. The vacuum system of conventional design was evacuated for a minimum of 10 hours at pressures less than 0.001 mm Hg. The system then remained static for several hours, and a change in pressure no larger than 0.001 mm Hg was an indication of sufficient outgassing.

Quenching vapor was then admitted to the counter manifold system to the desired pressure, which ranged from 2 to 20 mm Hg. Spectroscopically pure argon gas was introduced into the counters until the desired total pressure was attained. The counters were not used for about an hour after being filled, allowing quenching and vehicle gases to mix uniformly throughout the volume of the counters.

The high-voltage supply design given by Strong served as the basis for the circuit employed.[6] His circuit and component values were modified slightly to produce a regulated, stabilized voltage of 2000 volts with a drift of less than half a volt per hour. The error in voltage reading was less than 0.1 percent.

Spurious pulses in the breakdown region were of random height, mostly smaller than the primary pulse, and occurred at randomly rapid rates. Since the time interval between pulses was an important factor in counting all pulses, an RCA interval timer capable of counting pulses separated by as little as 1 microsecond was used. An ordinary preamplifier and pulse-shaping circuit converted the Geiger counter pulses to the required pulse height and rise time necessary to trigger the interval timer.

The number of ions per pulse could not be measured directly; therefore the average number of ions per pulse was measured instead. The simplest method was to store charge equal to the charge of the ions as they were collected at the cathode. This was performed by a bank of condensers placed between the cathode of the counter and ground. A bank of condensers totalling 100 microfarads was used inasmuch as it was not desirable to have a voltage drop across the condensers greater than 0.5 volt.

The system of switches, schematized as part of the complete operating circuit in Fig. 3, was designed to synchronize the time interval, pulses counted, and charge collected for a given period of data collection. The systems were coordinated to act simultaneously with an error of less than 0.01 second. The coordinated apparatus were the register, the stopwatch, the condenser bank, and an auxiliary current meter in parallel with the condenser bank. Four switches were operated simultaneously by one master lever. An electric stopwatch recorded the total time interval to within a tenth of a second.

Termination of the breakdown region was signified by three factors: (a) extremely rapid counting at a continuous rate even with the radioactive source removed; (b) arcing inside the counter between the electrodes, visible if the laboratory was darkened; and (c) an increasing current through the 5-μa microammeter driving the needle of the meter off scale.

Four types of data were tabulated: (a) counts

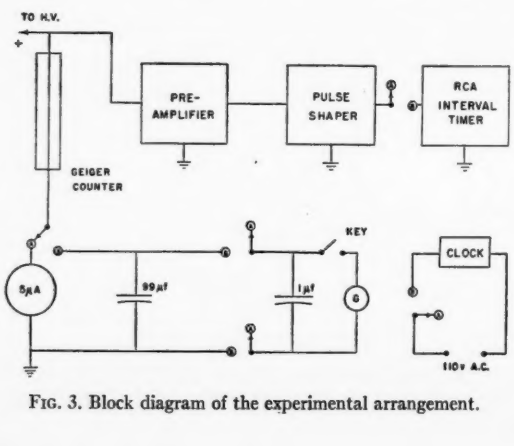


FIG. 3. Block diagram of the experimental arrangement.

[6] J. Strong, *Procedures in Experimental Physics* (Prentice-Hall, Inc., New York, 1946), p. 296.

recorded; (b) time of counting interval; (c) total charge collected in the counting interval; and (d) applied voltage.

From these data, plots were made of counting rate *versus* voltage and average number of ions per pulse *versus* voltage.

III. RESULTS

(a) Counting Rate *versus* Voltage

The characteristic plot of Geiger counter operation, counting rate *versus* voltage, is shown in Fig. 4. This plot established the plateau region in which the counting rate was constant with increasing voltage, or had a slight slope, and within which the Geiger counter normally would be operated for counting purposes. This experiment, however, was concerned with the high-voltage end of the graph where the counting rate increased sharply with small increases in voltage.

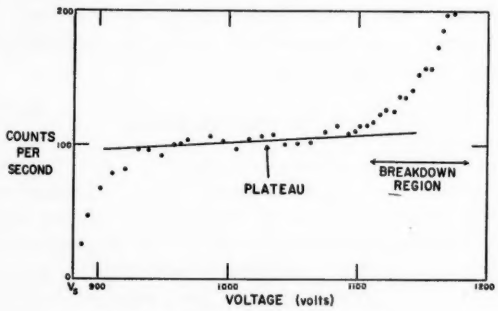


FIG. 4. Total counting rate *vs* voltage.

A straight line drawn through the plateau region gave the average counting rate. The rate of spurious counts was then defined as the difference between the total counting rate and the counting rate on the plateau.

(b) Charge per Pulse *versus* Voltage

A plot of the average number of ions per pulse *versus* voltage, shown in Fig. 5, was composed of two straight-line sections plus a portion which deviated from a straight line at the high-voltage end. Section I rose from zero ions per pulse at V_s to a point designated on the voltage scale as V_B. Section II then continued the straight-line rise from V_B, but with half the slope of Sec. I. This behavior was expected, having been explained by Wilkinson.[7] V_s is defined as the starting voltage, and may also be determined by passing a straight line through the proportional region in Fig. 4, the intercept on the voltage axis for zero counting rate giving V_s. V_B is the voltage at which the charge developed per pulse is equal to the charge across the counter necessary to bring the counter up to the applied voltage.

[7] D. H. Wilkinson, *Ionization Chambers and Counters* (Cambridge University Press, Cambridge, 1950), p. 182.

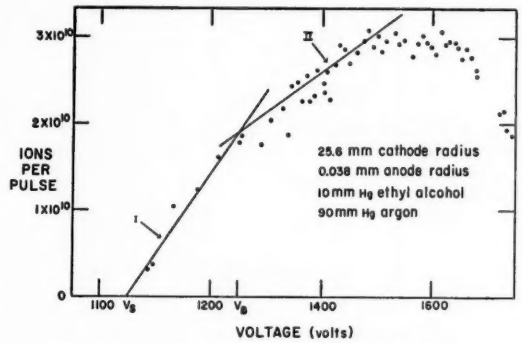


FIG. 5. Average number of ions per pulse *vs* voltage.

A deviation from the straight line of Sec. II was seen at the high-voltage end of Fig. 5, the experimentally determined points dropping toward values of lower ions per pulse. This implied that in this region although secondary pulses contributed to the total counting rate, some were too small to add appreciably to the total charge collected, thereby decreasing the average number of ions per pulse.

IV. DISCUSSION OF RESULTS

(a) Poisitive-Ion Bombardment

To determine whether the rate of spurious counts was related to the plateau counting rate, their ratio was plotted *versus* voltage (Fig. 6). Of course this ratio was zero on the plateau because, within the standard deviation, no spurious counts were detectable in this region. The graph was a straight line over a voltage range depending on the Geiger counter parameters, with a slope that varied from counter to counter. At the high-voltage end of the graph, the straight line changed rapidly into an exponentially rising curve.

The theory of positive-ion bombardment as a source of spurious counts in Geiger counters, as presented by Korff and Present, depended on the probability of a positive ion liberating a secondary electron at the cathode.[1] The ratio, secondary counts per positive ion,

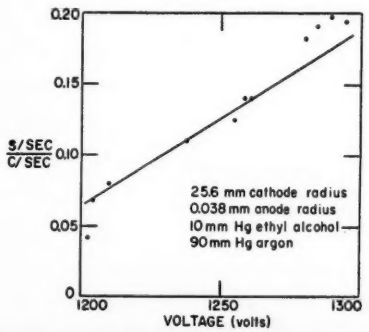


FIG. 6. Plot of the ratio of the number of spurious counts per second to the number of plateau counts per second *vs* voltage.

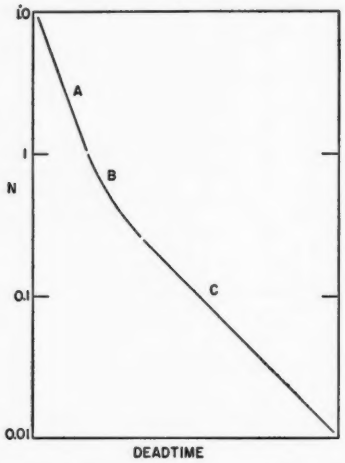


FIG. 7. Log$N$ *vs* deadtime, a composite graph.

was obtained by dividing the slope of Fig. 6 by the slope of Sec. II, Fig. 5, because the ratio secondary counts per plateau count was linear with voltage over the region considered, and ions per count was also linear with voltage over the same region. This ratio, secondary counts per positive ion, gave values in the range 1.8×10^{-11} to 5×10^{-11} secondary counts per ion, averaging 2.5×10^{-11} secondary counts per ion for all pressures and electrode diameters used, for argon and ethyl alcohol (since the ions bombarding the cathode are believed to be alcohol ions), and for brass cathodes. This agrees fairly well with the value 3.3×10^{-11} reported by Beretta and Rostagni for the same gases on brass cathodes.[8] It also agrees to order of magnitude with the value given by Korff and Present.[1]

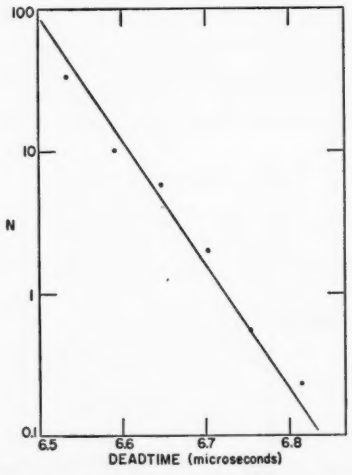


FIG. 8. Log$N$ *vs* deadtime for a total pressure of 40 mm Hg.

[8] E. Beretta and A. Rostagni, Nuovo cimento 6, 391 (1949).

The deviation, in Fig. 5, of the experimental points from the theoretically expected straight line at the high-voltage end of the plot was correlated with the onset of secondary counts at the beginning of the breakdown region in Fig. 4. The deviation in Fig. 5 did not occur exactly at the voltage at which secondary counts were first noticed in Fig. 4, but at a higher voltage. This voltage difference can be interpreted in the light of the fact that the graph of Fig. 6 was linear over a voltage region corresponding to the beginning of the breakdown region, implying, therefore, that the secondary pulses observed in this region are full-size pulses. Because of the continuation of straight line II over this voltage region, the pulses must have originated near or at the cathode since, only in this way, is a secondary electron capable of traversing the full electric field between the electrodes, giving rise to the full possible amplification on its flight to the anode and therefore to a full-size pulse. Thus, the term "secondary counts per plateau count" is more meaningful when expressed as secondary electrons emitted from the cathode per bombarding positive ion.

(b) Metastable Decay

Other results of this experiment indicate, however, that positive-ion bombardment was not the *dominant* responsible process, for it could not explain the rapid exponential rise in counting rate or in spurious counts per plateau count at the high-voltage end of the graphs. Some other cause must have been responsible for the remaining secondary counts. That these secondary counts were not of full size was illustrated in Fig. 5, where the number of ions per pulse decreased sharply at high voltages though the counting rate and total charge collected per second continued to increase. Although many secondary pulses were added to the total counting rate, they were too small to add appreciably to the charge collected, thereby decreasing the average number of ions per pulse. Furthermore, the rapid deviation from the straight line at high voltages in Fig. 6 implied that there were secondary counts originating in a region other than that near the cathode. Finally, a view of the pulses on a cathode-ray oscilloscope showed pulses in the breakdown region of various heights. Pulses of about 0.1 volt were seen, as small as could be detected by the counting apparatus. The height of pulses on the plateau was in the neighborhood of 7–10 volts.

These clues suggested that pulses were formed between the deadtime and recovery time of the Geiger counter, the smallest pulses having been formed immediately after the deadtime period, and the full-size secondaries from the cathode having been initiated after the recovery time. The small pulses formed close to the end of the deadtime period must have originated somewhere in the gas near the wire because the positive-ion sheath was still traveling across the counter at this time. Therefore, the possible role of metastable atom

and/or molecule decay through collision with other gas particles was investigated as a cause of these spurious counts.

Let us define N as the ratio of spurious counts (nonfull-size) to full-size pulses. The nonfull-size spurious counts will be shown to arise from metastable atom and molecule decay. The full-size pulses include both primary pulses and spurious pulses initiated by positive-ion bombardment of the cathode. Figure 7 is a composite plot of log$N$ *versus* the deadtime of the counter. The deadtime varied roughly as the reciprocal of the voltage, and it was calculated by using the formula given by Wilkinson whereby the deadtime is a function of the applied voltage V, V_s, V_B, anode radius, cathode radius, total pressure of the gases, and ion mobility[9]:

$$t_d=\frac{b^2\ln(b/a)\exp\left[\dfrac{-2(V-V_s)}{mV}\log b/a\right]}{2K(P/P_0)V[1+\frac{1}{2}m]}, \qquad (1)$$

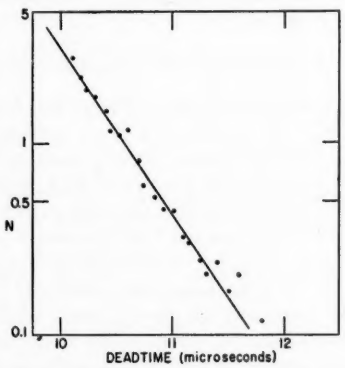


FIG. 9. Log$N$ *vs* deadtime for a total pressure of 200 mm Hg.

where m is equal to the ratio of the charge generated by the discharge per unit length of the wire to the charge per unit length of the counter required to charge it to its operating potential V.

Figure 7 is composed, on the whole, of two straight-line sections, A and C, connected by transition curve B. The slope of Sec. A is approximately four times the slope of Sec. C for argon-ethyl alcohol mixtures, whenever both are present in one graph. For low total pressures of 40 mm Hg (argon plus 10 percent alcohol), only the straight line A is present (Fig. 8). For high total pressures of 200 mm Hg, line C is the highly dominant portion (Fig. 9). At intervening pressures, 60 and 100 mm Hg, the transition curve B dominates the plot with traces of one or both straight-line portions, depending on counter parameters (Figs. 10 and 11).

The resultant theory, developed in this section, makes use of Meyerott's hypothesis that the proportion of metastable molecules to metastable atoms increases

[9] D. H. Wilkinson, see reference 7, p. 206.

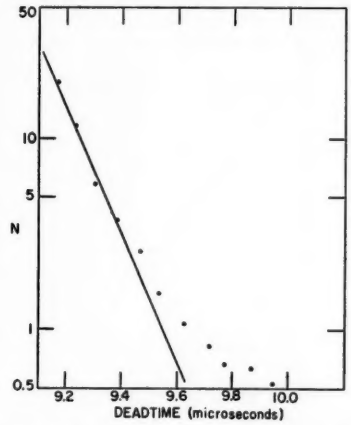


FIG. 10. Log$N$ *vs* deadtime for a total pressure of 60 mm Hg.

with the square of the pressure.[10] Inasmuch as metastable molecules are formed from metastable atom collisions, straight line A, dominant at low pressures, may be considered to represent collisions and subsequent secondary counts initiated by metastable atom decay; and straight line C, dominant at high pressures, may be considered to represent metastable molecule collision and decay. The term N, previously defined, is, of course, not the number of metastables decaying, but it does indicate that metastables existed and subsequently decayed after the deadtime of the initial primary pulse. The following theory shows the dependence of N on time, where time is measured from the initiation of the primary pulse as time zero which, within an infinitesimal time difference, is also the time of the formation of the large number of metastable atoms formed in the initial primary pulse.

Assume that M_0 metastable atoms were formed at time zero. Kruithof reported the ratio of metastable

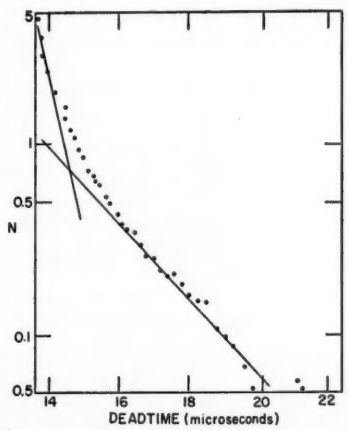


FIG. 11. Log$N$ *vs* deadtime for a total pressure of 100 mm Hg.

[10] R. Meyerott, Phys. Rev. **70**. 671 (1946).

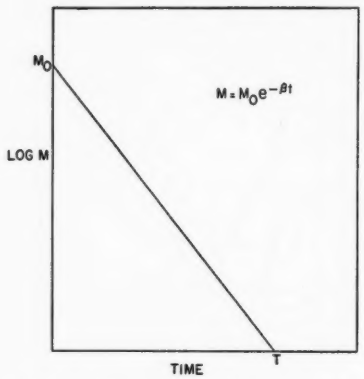


FIG. 12. Metastable atom decay curve.

atoms to ions as approximately 1:1 for argon in the E/p range employed in this research.[11] Because only the number of positive ions could be measured experimentally, this was the only available method of estimating M_0. At any subsequent time t, the number of metastable atoms existing, M, is given by

$$M=M_0e^{-\beta t}, \tag{2}$$

where β is the decay constant, the reciprocal of the mean life of the metastable atoms. This equation is shown plotted in Fig. 12.

The intercept on the time axis is the time T, statistically speaking, when the last metastable exists. When the deadtime of the primary pulse is larger than T, it can be expected, on the average, that no spurious pulse will be initiated. If the deadtime is decreased by increasing the applied voltage, until it coincides with T, then, on the average, one metastable exists that will initiate a spurious pulse, i.e., at this value of the deadtime, $N=1$. If the deadtime is decreased further, becoming less than T, then many metastables exist that may initiate spurious pulses. Only the metastables existing after the deadtime may initiate spurious pulses. However, not all metastables existing after the deadtime may do so, because some may decay during the deadtime of subsequent spurious pulses. M_0, the number of metastables per pulse, also shows an increase, although a slow one, with increasing voltage (Fig. 5). Figure 12 involves $\log M_0$, however; therefore, the increase in $\log M_0$ and in T, with increasing voltage, is very slow and not too important to this theory.

For metastable atoms, at the pressures used in this work, β is given in the form[12,13]

$$\beta=Ap+Bp^2+Cp_i, \tag{3}$$

where Ap represents decay through two-body collision with neutral vehicle gas atoms, Bp^2 represents decay through three-body collision with two neutral vehicle gas atoms, resulting in the formation of metastable molecules, and Cp_i represents decay through two-body collision with quenching gas molecules. The terms in Eq. (3) giving rise to spurious pulses through loss of excess energy by the metastable atom, probably by photon emission, are Ap and Cp_i. The quenching gas molecules, excited by collisions with metastable atoms, have a strong tendency to dissociate rather than reradiate, the fraction of excited molecules reradiating being only about 1 in 10^6. The rate of spurious counts per full-size pulse initiated by metastable atom decay is given, therefore, as

$$\frac{dN'}{dt}=10^{-5}(Ap+10^{-6}Cp_i)M_0e^{-\beta t}, \tag{4}$$

where $N=N'+N''$, N' is the portion of N initiated by metastable atom decay, N'' is the portion of N initiated by metastable molecule decay, and the coefficient 10^{-5} is the efficiency for electron emission by a photon striking the cathode. Integration of this equation gives N' after a time t as

$$N'=10^{-5}(Ap+10^{-6}Cp_i)M_0e^{-\beta t}/\beta. \tag{5}$$

For argon metastable atoms, Phelps and Molnar gave values of $A=40$/mm-sec and $B=9$/mm^2-sec.[12] These results for pure argon have recently been substantiated in this laboratory by A. Futch and F. A. Grant.

Using Eq. (3) for β in Eq. (5), and setting $M_0=10^{10}$, $N=1$, and $t=$deadtime$=9.1$ microseconds (t obtained from Fig. 10 at the point on the curve where $N=1$), the value 1×10^5/mm-sec for C was obtained. This result shows the dominance of the Cp_i term in Eq. (3) although, at 200 mm Hg pressure, Bp^2 may attain values of about 10 percent of Cp_i. It also indicates that Ap is dominant over 10^{-6} Cp_i in Eq. (4) even for 20 percent quenching vapor, the largest used.

Equation (4) can now be rewritten as

$$N'=\frac{10^{-5}ApM_0}{Cp_i}\exp[-(Bp^2+Cp_i)t]. \tag{6}$$

The term Bp^2 was retained in the exponent because it may have a value as much as 10 percent of Cp_i—of possible significance in an exponent.

The formation of metastable molecules in the decay of metastable atoms is expressed by the Bp^2 term. The dominant decay of the metastable molecules is believed to be both by spontaneous decay and by collision with quenching gas molecules. Colli and Facchini reported a spontaneous-decay mean life of 3 microseconds for pure argon.[3] Therefore the rate of change of the number of metastable molecules with time can be expressed as

$$dM''/dt=Bp^2M-(k+Dp_i)M'', \tag{7}$$

where M'' is the number of metastable molecules existing at any time t, k is the reciprocal of the spontaneous-

[11] A. A. Kruithof, Physica 7, 519 (1950).
[12] A. V. Phelps and J. P. Molnar, Phys. Rev. 89, 1202 (1953).
[13] M. A. Biondi, Phys. Rev. 82, 453 (1951).

decay mean life, and D is a coefficient representing the number of collisions with quenching molecules leading to decay. Substituting for M from Eq. (2) and integrating, gives

$$M''=\frac{Bp^2M_0}{Cp_i-(k+Dp_i)}\{\exp[-(k+Dp_i)t]-\exp[-(Bp^2+Cp_i)t]\}. \quad (8)$$

Here again Bp^2 in conjunction with Cp_i is retained only in the exponent.

N'' is the ratio of spurious pulses to full-size primary pulses where the spurious pulses are initiated by metastable molecule decay after a time t. The decay term in Eq. (7) is $(k+Dp_i)M''$; therefore, with the fraction of quenching molecules reradiating equal to 1 in 10^6, and with the photon efficiency equal to 10^{-5},

$$dN''/dt=10^{-5}(k+10^{-6}Dp_i)M''. \quad (9)$$

Inserting Eq. (8) for M'', neglecting the term $10^{-6}\,Dp_i$ in comparison with k, and integrating, we find

$$N''=\frac{10^{-5}kBp^2M_0}{Cp_i-(k+Dp_i)}\left[\frac{\exp[-(k+Dp_i)t]}{k+Dp_i}-\frac{\exp[-(Bp^2+Cp_i)]}{Cp_i}\right]. \quad (10)$$

The graphs of Fig. 7 that contained some portions of both Secs. A and C, for many combinations of gas pressures, gave a ratio of the slope of A to the slope of C of about 4:1 for argon-ethyl alcohol mixtures. Since Fig. 7 was a semilogarithmic plot, this ratio of the slopes was interpreted as being the ratio of the coefficients (Bp^2+Cp_i) to $(k+Dp_i)$ in the exponents of the equations. This signifies that the mean life of metastable molecules is approximately 4 times the mean life of metastable atoms with ethyl alcohol as the quenching agent in the range of gas pressures used. Therefore the term

$$\exp[-(k+Dp_i)t]/(k+Dp_i)$$

is dominant over

$$\exp[-(Bp^2+Cp_i)t]/Cp_i$$

for the pressures and deadtimes involved in the experiments. Therefore,

$$N''=\frac{10^{-5}kBp^2M_0}{Cp_i-(k+Dp_i)}\left[\frac{\exp[-(k+Dp_i)t]}{k+Dp_i}\right]. \quad (11)$$

Solving Eq. (11), setting $M_0=10^{10}$, $N''=1$, and t=deadtime=25 microseconds (t obtained from Fig. 7, for 5 mm Hg alcohol and 200 mm total pressure, at the point on the curve where $N=1$), the value 2×10^4/mm-sec for D was obtained. This means that ethyl alcohol as a quenching agent is approximately 5 times as effective on argon metastable atoms as on argon metastable molecules, since $C=5D$. Furthermore, $k=3\times10^5$ indicates that spontaneous metastable molecule decay is of the same order of magnitude as Dp_i, the decay through collisions with the quenching vapor molecules.

Experimental results using Eqs. (5) and (11) are tabulated in Table I for counters with 7.14 mm cathode radius and in Table II for counters with 25.6 mm cathode radius, for argon and ethyl alcohol mixtures. At the dead time, the expected value for N is unity, since when $T=t_d$ (Fig. 12) only one spurious count is expected to be initiated by metastable decay. That most of the results are of the same order of magnitude as unity, is considered good agreement because factors such as the efficiency of a photon liberating a secondary electron at the cathode, and the proportion of ethyl

TABLE I. Summary of results, argon-ethyl alcohol, 7.14 mm cathode radius.

Total pressure (mm Hg) p	Quenching gas pressure (mm Hg) p_i	Deadtime (microseconds) t (for $N=1$)	N'	N''
20	2	3	100	
40	4	6.7	15	
		6.7	15	
		6.8	15	
60	6	8.5	2	25
		9.1	1	20
		9.3	1	20
100	5	20	0.1	10
		27	0.2	0.3
100	10	11	0.01	15
		14	0.001	3
100	20	8		3
200	5	19		15
		24		3
200	10	13		15
		14		10
200	20	10		3
		11		2
		15		0.5

TABLE II. Summary of results, argon-ethyl alcohol, 25.6 mm cathode radius.

Total pressure (mm Hg) p	Quenching gas pressure (mm Hg) p_i	Deadtime (microseconds) t (for $N=1$)	N''
100	5	22	2
		24	1
		25	0.5
100	10	14	0.3
		15	0.1
200	5	21	1
		25	1

alcohol excited molecules reradiating are known only to an order of magnitude.

(c) Argon-Ethyl Acetate

Argon-ethyl acetate results at 40 mm Hg and 200 mm Hg total pressures, with 10 percent quenching vapor, show longer plateaus that those for the argon-ethyl alcohol mixtures. For example, at 40 mm Hg, the ethyl alcohol plateau is only about 5 volts wide, whereas the acetate plateau is about 40 volts wide. The value of C, determined for ethyl acetate from the preceding equations, was about 1.1×10^5, or approximately the same as C for ethyl alcohol, where C was 1×10^5. To explain the difference in plateau widths obtained between ethyl alcohol and ethyl acetate, a difference of only 10 percent in the value of C is required. However, this is within the experimental error, and, therefore, further conclusions cannot be made. The value of D was the same for ethyl acetate as for ethyl alcohol, within the experimental error.

ACKNOWLEDGMENTS

The authors wish to thank Professor F. A. Grant for much valuable information concerning metastable atoms and Professor R. D. Myers for his many helpful discussions during the course of this work.

PHYSICAL REVIEW VOLUME 98, NUMBER 2 APRIL 15, 1955

High-Frequency Gas Discharge Plasma in Hydrogen*

DAVID J. ROSE, *Bell Telephone Laboratories, Murray Hill, New Jersey*

AND

SANBORN C. BROWN, *Department of Physics and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts*

(Received January 13, 1955)

The high-frequency electric field required to maintain a hydrogen plasma has been measured as a function of pressure and plasma electron density. A theory of the plasma based on a solution of the Boltzmann transport equation has been developed to predict this field; it agrees satisfactorily with experiment. The theory has no adjustable parameters, and uses only the probabilities of collision, excitation, and ionization of the gas by electrons, and the ionic mobility.

I. INTRODUCTION

MACDONALD and Brown[1] developed a theory of the mechanism of breakdown in hydrogen at high frequencies that is based on a solution of the Boltzmann transport equation. Calculated breakdown fields agree well with their experimental results. More recently, Allis and Brown[2] (A-B) simplified the method of solution and calculation of the high-pressure case, in which electrons diffusing from the discharge do not remove an appreciable fraction of the total energy. This simplified method may also be applied to the calculation of breakdown in certain other gases. The present paper generalizes these theories by taking into account the effect of space charge, so that it will be possible to compute from first principles the microwave field strength required to maintain certain simple plasmas in the steady state. Calculations made on this basis will be compared with experimental measurements obtained for hydrogen.

The theory presupposes that the following additional nonlinear phenomena peculiar to some plasmas and not to breakdown may be neglected: (a) cumulative ionization, (b) electron-ion recombination, (c) formation of negative ions in appreciable numbers, (d) electron-electron interactions, (e) plasma resonance. These restrictions place an upper limit on the plasma electron density for which the theory is applicable. Limits arising from (a), (b), and (c) may be readily calculated. Regarding limit (d), the work of Haseltine[3] shows that the electron-electron interaction has least effect in gases in which the collision frequency of electrons is independent of electron energy. *A posteriori* calculations show that for hydrogen limit (e) actually controls in our experiments, at frequencies of about 3000 Mc/sec, above a density limit of about $10^{11}/cm^3$, a value well above those experimentally obtained here.

The physical process is the following: Electrons gain energy from the applied field, and lose energy by elastic and inelastic collisions. Ionization of gas molecules provides a source of new electrons, and flow to the tube walls in the presence of density and space-charge potential gradients provides the sink. The Boltzmann transport equation in phase space expresses these balances both in energy and in number; correspond-

* This work was supported in part by the Signal Corps; the Office of Scientific Research, Air Research and Development Command; and the Office of Naval Research.

[1] A. D. MacDonald and S. C. Brown, Phys. Rev. **76**, 1634 (1949).

[2] W. P. Allis and S. C. Brown, Phys. Rev. **87**, 419 (1952).

[3] W. R. Haseltine, J. Math. Phys. **18**, 174 (1939).

ingly, a solution is obtained by splitting it into two parts.

The energy part will contain certain constant spatial parameters describing the space-charge field and diffusion length, and will yield the electron velocity distribution. The spatial part will contain other constant energy parameters (electron diffusion, mobility, and ionization rate), and will give the current and space distributions. The complete solution is obtained by making these various parameters self-consistent.

The problem of the spatial distributions and current flow have been treated by Allis and Rose[4] (A-R) and these results will be used here. The effect of the space-charge field upon the energy distribution, from which the ionization rate is to be computed, must now be considered.

II. THE BOLTZMANN EQUATION

The development of the Boltzmann equation will follow the work of A-B as closely as possible, and their notation will be used. The only difference is that a dc space-charge field E_s must be superimposed on the high-frequency field so that

$$\mathbf{E}=\mathbf{E}_s+\mathbf{E}_p\exp(j\omega t), \tag{1}$$

where E_s varies with position, but E_p is assumed constant. The distribution F of electrons in velocity and configuration space is then determined by the Boltzmann equation

$$\partial F/\partial t=C-\nabla_r\cdot\mathbf{v}F+\nabla_v\cdot e\mathbf{E}F/m. \tag{2}$$

Here, C represents the effect of collisions; ∇_r and ∇_v are the gradient operators in configuration and velocity space; and v, e, and m are the velocity, charge, and mass of the electron. Following A-B, F is expanded in spherical harmonics in velocity space and a Fourier series in time:

$$\begin{aligned}F&=\sum_l\sum_k F_k^l P_l(\cos\theta)\exp(jk\omega t),\\&=F_0^0+\mathbf{v}\cdot[\mathbf{F}_0'+\mathbf{F}_1'\exp(j\omega t)]/v.\end{aligned} \tag{3}$$

It may be shown that the three indicated items suffice for the expansion discussed by Allis and Brown, even if space charge is present, provided that the electron density is sufficiently low that plasma resonance is not encountered.

Allis and Brown have evaluated the collision terms, and have separated the various harmonic terms of the distribution. There are one scalar and two vector equations:

$$\begin{aligned}(\nu_x+\nu_i-q)F_0^0=&-(v/3)\nabla_r\cdot\mathbf{F}_0'+(1/v)^2\partial\{[ev^2/6m]\\&\times[(\mathbf{E}_p\cdot\mathbf{F}_1')_{\text{real}}+2\mathbf{E}_s\cdot\mathbf{F}_0']\\&+[m/M]\nu_c v^3F_0^0\}/\partial v,\end{aligned} \tag{4}$$

$$\nu_c\mathbf{F}_0'=-v\nabla_r F_0^0+(e/m)\mathbf{E}_s\partial F_0^0/\partial v, \tag{5}$$

$$(\nu_c+j\omega)\mathbf{F}_1'=(e\mathbf{E}_p/m)\partial F_0^0/\partial v. \tag{6}$$

[4] W. P. Allis and D. J. Rose, Phys. Rev. **93**, 84 (1954).

Here, ν_c, ν_x, and ν_i are the collision frequencies of an electron for momentum transfer, excitation, and ionization, respectively. The quantity qF_0^0 is the rate of appearance of new electrons at low energies as a result of excitation and ionization; it is permissible to set it equal to a delta function at the origin of velocity space. Except for the additional field $\mathbf{E}_s$, these equations are identical with Eqs. (5), (6), and (7) of A-B.

III. SEPARATION OF F IN SPACE AND ENERGY

A solution of Eqs. (4)–(6) is possible if F may be separated into spatial and energy parts, $F=N(\mathbf{r})f(v)$. This means, in effect, that the velocity distribution must be the same everywhere in the discharge, and therefore that the diffusion coefficient and mobility are constant.

It is possible to obtain the set of spatial equations by integrating Eqs. (4) and (5) over velocity space. The dc electron particle current,

$$\mathbf{\Gamma}=\int_0^\infty\mathbf{F}_0'(4\pi v^3/3)dv=-\nabla_r(D_-N)-\mu_-\mathbf{E}_sN, \tag{7}$$

is determined by $\mathbf{F}_0'$ from Eq. (5). Here N is the electron spatial density, and the diffusion and mobility coefficients are

$$D_-N=\int_0^\infty(v^2/3\nu_c)F_0^0 4\pi v^2dv; \tag{8}$$

$$\mu_-N=\int_0^\infty(4\pi/3)(e/m)F_0^0[\partial(v^3/\nu_c)/\partial v]dv. \tag{9}$$

The total excitation and ionization rates are

$$N\langle\nu_x\rangle=\int_0^\infty\nu_xF_0^0 4\pi v^2dv, \tag{10}$$

$$N\langle\nu_i\rangle=\int_0^\infty\nu_iF_0^0 4\pi v^2dv; \tag{11}$$

also

$$N(\langle\nu_x\rangle+2\langle\nu_i\rangle)=\int_0^\infty qF_0^0 4\pi v^2dv \tag{12}$$

represents the appearance of all new electrons. Equation (4) may be multiplied by $4\pi v^2dv$ and integrated over all velocities. The term in braces vanishes at both limits, leaving

$$N\langle\nu_i\rangle=\nabla_r\cdot\mathbf{\Gamma}. \tag{13}$$

There is, in addition, an equation for the ion particle current which in the steady state also equals $\mathbf{\Gamma}$:

$$\mathbf{\Gamma}=-\nabla_r(D_+P)+\mu_+\mathbf{E}_sP, \tag{14}$$

where P is the ion density. Finally, Poisson's equation

$$\nabla_r\cdot\mathbf{E}_s=e(P-N)/\epsilon_0 \tag{15}$$

must hold. Equations (7), (13)–(15) determine the spatial distributions from which the average ionization frequency per electron $\langle\nu_i\rangle$ is determined as a characteristic value. It is these equations, with constant D's and μ's, that have been treated by Allis and Rose.

The presence of certain crossed space and velocity terms in Eqs. (4)–(6) indicate that the separation of F into a product $N(r)F(v)$ is not quite correct. It will now be shown, however, that the dependence of $f(v)$ on space is weak, so that the separation is a good approximation. First, the magnitudes of the terms $(\mathbf{E}_p\cdot\mathbf{F}_1')_{\text{real}}$ and $2\mathbf{E}_s\cdot\mathbf{F}_0'$ are compared. The first represents the energy gain of the electrons from the applied alternating field, and the second their energy loss in flowing against the space-charge field. At values of $p\Lambda>1$ each electron gains energy from E_p sufficient for many exciting collisions and one ionizing collision, in addition to that required to overcome the elastic recoil losses. This total energy is many times the wall potential, so that the space-charge term in Eq. (4) is negligible. The $\mathbf{E}_s$ term in Eq. (5), however, cannot be dismissed, for it represents an inward mobility flow that almost cancels the outward diffusion term.

Equations (5) and (6) are substituted in Eq. (4), giving

$$\begin{aligned}(\nu_x+\nu_i-q)F_0^0=&(v^2/3\nu_c)\nabla_r^2F_0^0\\&-(ev/3m\nu_c)\nabla_r\cdot(\mathbf{E}_s\partial F_0^0/\partial v)\\&+(1/v)^2\partial[(eu_c/3m)\nu_cv^2(\partial F_0^0/\partial v)\\&\qquad+m\nu_cv^3F_0^0/M]/\partial v,\end{aligned}\quad(16)$$

where

$$u_c=eE_p^2/2m(\nu_c^2+\omega^2).\quad(17)$$

The terms of Eq. (16) represent, in order, the excitation and ionization processes, diffusion, the counteracting space-charge mobility, energy gain from the applied field, and elastic recoil.

In cases where the electron density is low, $\mathbf{E}_s$ is small, and only the diffusion term contains spatial derivatives. Equation (16) then separates with $\nabla_r^2F_0^0=-F_0^0/\Lambda^2$, where Λ is the cavity diffusion length. The spatial distribution is determined by the diffusion equation. On the other hand, when N and $\mathbf{E}_s$ are large, the mobility and diffusion terms nearly cancel, and their difference is small compared to the other terms in the equation; an approximation for the small difference will therefore suffice. A separation is still possible provided only that $\mathbf{E}_sN\sim\nabla_rN$, in which case $\nabla_r^2N=-N/\Lambda^2$ as before. This solution of the space equations is the constant ratio approximation discussed by Allis and Rose. It is correct both at very low and very high electron densities, but is somewhat in error at intermediate densities.

We assume, then, that

$$\mathbf{E}_s=-u_s\nabla N/N,\quad(18)$$

where u_s is a measure of the space charge. We shall also define the electron energy $u=mv^2/2e$ in electron-volts; from Eq. (16), one then obtains

$$\begin{aligned}&(\nu_x+\nu_i-q)f+(2eu/3m\nu_c\Lambda^2)(f+u_sdf/du)\\&\quad=(2/3\sqrt{u})d[u^{\frac{1}{2}}\nu_c(u_cdf/du+3mf/M)]/du.\end{aligned}\quad(19)$$

Let an effective field E_e be defined as

$$E_e^2=\nu_c^2E_p^2/2(\nu_c^2+\omega^2).\quad(20)$$

When the specific dependences of ν_x, ν_i, and ν_c are inserted, Eq. (19) may be solved to yield a relation between the three parameters p/E_e, $E_e\Lambda$, and u_s. Furthermore, Eq. (19) may be multiplied by $4\pi v^2dv$ and integrated over all velocities. The right side vanishes, and from Eqs. (8)–(12)

$$D_--u_s\mu_-=\langle\nu_i\rangle\Lambda^2\equiv D_s,\quad(21)$$

which defines D_s, the effective diffusion coefficient. The escape frequency of electrons, measured by D_s/Λ^2, is reduced from the free diffusion frequency D_-/Λ^2 on account of the space-charge field measured by u_s; $\langle\nu_i\rangle$ is correspondingly reduced, and so is the field E_e which maintains the ionization. Thus p/E_e rises in the steady state over the breakdown value.

From the solution of Eqs. (19) and (21), D_-, μ_- and D_s/D_- may be computed as functions of p/E_e and $E_e\Lambda$. Since Allis and Rose give D_s/D_- as a function of $N_0\Lambda^2\mu_-/D_-$, where N_0 is the electron density at the center of a parallel plane cavity, it is possible to plot contours of constant $N_0\Lambda^2$ on the $p/E_e-E_e\Lambda$ plane. The special case $N_0\Lambda^2=0$, whence $D_s/D_-=1$ and $u_s=0$, corresponds to breakdown.

IV. SOLUTION OF THE ENERGY DISTRIBUTION

It is convenient to consider solutions of Eq. (19) for the case of large $p\Lambda$ (greater than 1 mm Hg×cm) and small $p\Lambda$ (between 0.1 and 1.0 mm Hg×cm) separately. All of the experimental measurements were taken within the limit of validity of diffusion theory.[5]

At large values of $p\Lambda$, hence also of $E_e\Lambda$, the escape frequency is low, and the effect of the current flow on the energy distribution becomes small. Correspondingly, the second term of Eq. (19), being in ratio $1/(p\Lambda)^2$ to the remaining terms, is small. In this range, f is determined principally by the balance between energy gain and elastic recoil loss, since there are few electrons in the inelastic range. If ν_c is independent of electron energy, as is approximately the case with hydrogen, f is close to Maxwellian. The small second term of Eq. (19) can then be approximated by replacing df/du by $-\mu_-f/D_-$, which is true in the average in any case. With the aid of Eq. (21), Eq. (19) becomes

$$\begin{aligned}&(\nu_x+\nu_i-q)f+(2eu/3m\nu_c\Lambda_e^2)f\\&\quad=(2/3\sqrt{u})d[u^{\frac{1}{2}}\nu_c(u_cdf/du+3mf/M)]du,\end{aligned}\quad(22)$$

where

$$\Lambda_e^2=\Lambda^2D_-/D_s\quad(23)$$

[5] S. C. Brown and A. D. MacDonald, Phys. Rev. **76**, 1629 (1949).

defines an effective diffusion length that is larger than the actual cavity diffusion length when the space-charge field is important. Equation (22) is the breakdown equation of A-B for a cavity of diffusion length Λ_e, and the existing methods of solution may be applied directly. Steady-state values of $E_e\Lambda$ *vs* p/E_e will be replicas of the breakdown values, shifted to lower $E_e\Lambda$ by the ratio $(D_e/D_-)^{\frac{1}{2}}$.

At lower values of $p\Lambda$, hence also of $E_e\Lambda$ and p/E_e, the diffusion term is not negligible, and simple approximations cannot be made for it. It is then necessary to develop solutions of the general sort derived in reference 1. This has been done for the case of constant ν_c. Define the parameter

$$\sigma=(2m/M)-(2eu_s/3m\nu_c^2\Lambda^2), \tag{24}$$

which is positive at large $p\Lambda$, and negative at small $p\Lambda$ and large u_s. Different solutions are required, depending on the sign of σ. They differ only in slight mathematical detail; thus, in the following, let $\sigma=|\sigma|$. Then the upper and lower signs, where applied, represent the positive and the negative cases, respectively. The special case $\sigma=0$ has no particular significance.

Define the additional parameters:

$$u_e=2u_c/3\sigma, \tag{25}$$

$$1/\beta^2=1+(2u_e/E_e\Lambda)^2, \tag{26}$$

$$Q=6(2m/\sigma M-1), \tag{27}$$

and a new energy variable

$$z=u/u_e\beta. \tag{28}$$

The quantity q in Eq. (19) will be set equal to a delta function at the origin, and the excitation plus ionization rate will be approximated by

$$\nu_x+\nu_i=\nu_c h_x(u-u_x), \tag{29}$$

where u_x is the first excitation potential. Equation (19) is now expressed in normal form in terms of the variable z, so that

$$f=z^{-\frac{1}{2}}g\exp(\mp\beta z/2), \tag{30}$$

$$d^2g/dz^2=[A^2+B/z-3/16z^2]g, \tag{31}$$

where

$$A^2=(\sigma+4h_xu_e\beta^2)/4\sigma, \tag{32}$$

$$B=[(Q\pm3)\sigma\beta+4h_xu_e\beta^2z_x]/4\sigma. \tag{33}$$

The body of the distribution function is that part lying below u_x; in that region $\nu_x=\nu_i=0$, and one must set $h_x=0$ in Eqs. (31)–(33). The solution in the body is then

$$f=\exp[-z(1\pm\beta)/2][-CM_1(z)+M_2(z)], \tag{34}$$

where

$$M_1=M\{[3-(Q\pm3)\beta]/4;3/2;z\}, \tag{35}$$

$$M_2=z^{-\frac{1}{2}}M\{[1-(Q\pm3)\beta]/4;1/2;z\}, \tag{36}$$

and $M(\alpha;\gamma;z)$ is the confluent hypergeometric function.[6] The constant C will be determined later.

The tail of the distribution lies above u_x, and the function is given by Eqs. (31)–(33) directly. Rather than solving exactly, it is much simpler to use the asymptotic expansion

$$g=\exp(-Az)z^b(1-d/z+\cdots). \tag{37}$$

The invariant I of Eq. (37) is

$$I=A^2-2Ab/z-(1/z)^2[b-b^2+2Ad/(1-d/z)-2bd/z(1-d/z)+2d/z(1-d/z)], \tag{38}$$

which is to be fitted as well as possible to the invariant of Eq. (31). Thus

$$b=B/2A, \tag{39}$$

satisfying the term in $1/z$. The remaining term in $1/z^2$ will be set equal to $\frac{3}{16}$ at $z=z_x$, so that

$$1/d=[2A-2(b-1)/z_x]/[b(b-1)+\tfrac{3}{16}]+1/z_x. \tag{40}$$

Above u_x, we have, then,

$$f=R[\exp(-Sz)]z^T(1-d/z), \tag{41}$$

where

$$S=A+(\beta/2), \tag{42}$$

$$T=b-\tfrac{3}{4}. \tag{43}$$

Here, R and C are determined by matching value and slope of f at z_x. The derivative $M_2'(z)$ can be eliminated by use of the Wronskian

$$W(z)=M_1'M_2-M_2'M_1=[\exp(z)]/2z^{\frac{1}{2}}. \tag{44}$$

Then

$$R=\frac{\exp[z_x(1-\beta)/2]\exp Sz_x}{2{z_x}^V(1-d/z_x)[M_1'(z_x)+KM_1(z_x)]}, \tag{45}$$

$$C=\{M_2'(z_x)-(\exp z_x)/2z_x[M_1'(z_x)+KM_1(z_x)]\}/M_1(z_x), \tag{46}$$

where

$$V=T+\tfrac{3}{2}, \tag{47}$$

$$K=A-\tfrac{1}{2}-[T+d/z_x(1-d/z_x)]/z_x. \tag{48}$$

The distribution function f is now determined. There are two methods at hand for completing the solution. One method consists, first, of computing D_- and μ_- from Eqs. (8) and (9), the integration being performed numerically.

One may set, similarly to Eq. (29),

$$\nu_i=\nu_c h_i(u-u_i), \tag{49}$$

where u_i is the ionization potential. Then

$$\langle\nu_i\rangle=\nu_c h_i\int_{u_i}^{\infty}(u-u_i)f4\pi v^2dv$$

$$=2\pi\nu_c h_i R(2e/m)^{\frac{1}{2}}(u_e\beta)^{5/2}\times\int_{z_i}^{\infty}[\exp(-Sz)]z^T(1-d/z)(z-z_i)dz. \tag{50}$$

[6] A. D. MacDonald, J. Math. Phys. **28**, 183 (1949).

Let

$$y=Sz. \tag{51}$$

The integral in Eq. (50) may be performed by parts and expressed in terms of the incomplete gamma function[7] $I(y_i; V)$. Then

$$\langle \nu_i \rangle = 2\pi \nu_c R J_i (2e/m)^{\frac{1}{2}} (u_e\beta)^{5/2}/S^{V+1}, \tag{52}$$

where

$$\begin{aligned} J_i = h_i\{&[1-Sd/V-y_i/V \\ &+Sdy_i/V(V-1)][1-I(y_i; V)]\Gamma(V+1) \\ &-(y_i{}^V/\exp y_i)[Sdy_i/V(V-1) \\ &-(y_i/V)+Sd/V(V-1)]\}. \end{aligned} \tag{53}$$

The values of D_-, μ_-, and $\langle \nu_i \rangle$ are now substituted in Eq. (21), which is a transcendental equation involving p/E_e, $E_e\Lambda$, and u_s. By this method D_-/μ_- is also determined; it can thus be demonstrated that D_-/μ_- is practically independent of u_s, and is a function of p/E_e alone, as is shown in reference 2.

Once this fact is established, a second and somewhat less laborious method may be used. It consists in evaluating Eq. (12), after insertion of f. The integral has a value at the lower limit, and one obtains

$$\langle \nu_{xi} \rangle + \langle \nu_i \rangle = \pi (2e/m)^{5/2} (u_e\beta)^{\frac{1}{2}} E_e{}^2/3\nu_c, \tag{54}$$

where $\langle \nu_{xi} \rangle = \langle \nu_x \rangle + \langle \nu_i \rangle$. In principle the two methods are related through integration by parts; but the approximation made in Eq. (37) will, in practice, introduce small and different errors. $\langle \nu_i \rangle$ is given in Eq. (52); $\langle \nu_x \rangle$ and J_x are identical with Eqs. (52) and (53), with the subscript x rather than i. Equation (54) becomes

$$2u_e\beta^2 R(J_x+J_i)/\sigma S^{V+1}=1. \tag{55}$$

This is again a transcendental equation in p/E_e, $E_e\Lambda$, and u_s. Through Eq. (21), and A-R, one obtains the desired relation between p/E_e, $E_e\Lambda$ and $N_0\Lambda^2$.

V. EXPERIMENTAL PROCEDURE

A block diagram of the experimental apparatus is shown in Fig. 1. The discharge cavity is constructed of OFHC copper. It is a right circular cylinder with a radius of 3.57 cm and a height of 0.635 cm that gives a resonant wavelength of 9.4 cm in the TM_{010} mode, and

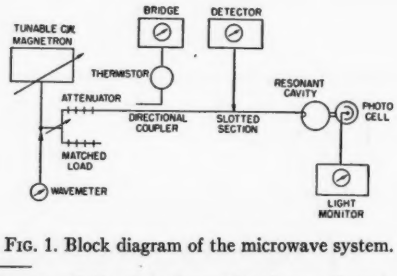


FIG. 1. Block diagram of the microwave system.

a diffusion length of 0.200 cm. Hydrogen can be admitted either through a heated palladium leak tube or directly from flasks of spectroscopically pure gas. A liquid nitrogen trap isolates the cavity and leak tube from all parts of the vacuum system which cannot be outgassed at temperatures above 300°C. The vacuum system can be pumped to pressures of the order of 10^{-8} mm Hg.

Experiments were performed in hydrogen in the pressure range 0.5–20 mm Hg. The procedure for each run was as follows: The gas was admitted to the desired pressure, and the discharge was initiated by adjusting the incident power level with the power divider. The light output of the discharge, monitored by a photocell, was noted. The magnetron frequency was then varied, the incident power being adjusted so that the light output was the same. At each frequency, the standing wave ratio and position of the voltage minimum on the slotted section, and the incident power, as read on the thermistor bridge, were recorded. In this way, a part of the resonance curve of the cavity plus discharge was determined. The electron density (and light output) was then set to a new value by changing input power, and the measurements were repeated. These measurements, plus the known characteristics of the empty cavity, suffice to determine the maintaining electric field and electron density.[8]

The measurements give both the discharge conductance g_d and susceptance b_d. The ratio is

$$\frac{g_d}{b_d} = \int_0^\infty \frac{u^{\frac{1}{2}}\nu_c}{(\nu_c{}^2+\omega^2)} \frac{df}{du} du \Big/ \int_0^\infty \frac{u^{\frac{1}{2}}\omega}{(\nu_c{}^2+\omega^2)} \frac{df}{du} du, \tag{56}$$

as can be seen from the integral over the velocities of Eq. (6). This gives the ac current which, in turn, is directly related to the complex conductivity of the plasma through Eq. (8) of Part III of reference 8. The right side of Eq. (56) is a weighted average $\langle \nu_c \rangle/\omega$. The present measurements yield, for hydrogen,

$$\langle \nu_c \rangle = 4.85\times 10^9 p \text{ sec}^{-1} \pm 2 \text{ percent}.$$

The value is constant over the range $0.01 < p/E_e < 0.10$ cm×mm Hg/volt covered in this experiment. The probability of collision $P_c=\nu_c/pv$ for hydrogen, above 4 ev is proportional to $1/v$, and in that region $\nu_c \approx 5.9\times 10^9 p$; below this value it is approximately constant. If the actual value of $\nu_c(v)$ is averaged over the distribution functions of Sec. IV, a value close to $4.85\times 10^9 p$ results for all values of E_e/p experimentally realized.

The experimental results, obtained in a TM_{010} mode cavity, do not correspond exactly to results that would be obtained in a plane-parallel system of infinite extent. Two corrections must be made. The first takes into account the radial variation of the applied field. The field varies with radius r as $J_0(2.405r/R)$, where J_0 is

[7] K. Pearson, *Table of the Incomplete Gamma Function* (Cambridge University Press, Cambridge, 1946).

[8] S. C. Brown and D. J. Rose, J. Appl. Phys. **23**, 711, 719, 1028 (1952).

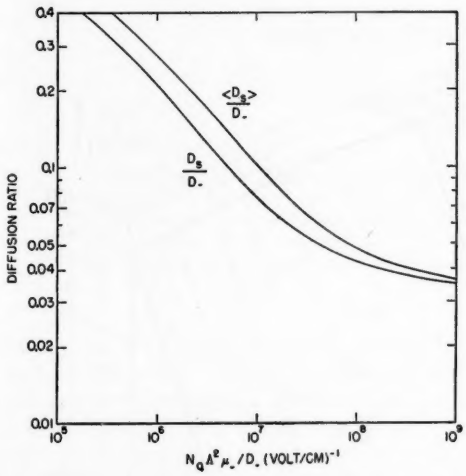


FIG. 2. Theoretical calculation of the quantities D_s/D_- and $\langle D_s\rangle/D_-$ as a function of $N_0\Lambda^2\mu_-/D_-$.

the zero-order Bessel function, and R is the cavity radius. The measured field at the center must be decreased slightly to correspond to the equivalent constant field throughout the cavity. The data are therefore corrected according to the theory of Herlin and Brown,[9] the correction amounting to no more than 5 percent.

The second correction arises because the plasma density itself has a radial variation, and this introduces a variable D_s for diffusion to the ends of the cylinder. If the field were uniform, a radial dependence of N and P,

$$\begin{Bmatrix} N \\ P \end{Bmatrix} = \begin{Bmatrix} N_0 \\ P_0 \end{Bmatrix} J_0(2.405r/R), \tag{57}$$

would be correct at both the free and ambipolar limits, and approximately correct in the transition region. Since the height of the cavity is much less than its radius, it is possible to define an average $\langle D_s\rangle$ for the discharge. It is

$$\langle D_s\rangle = \frac{\text{(Total electron flow to the walls)}}{\text{(Total number of electrons present)}} \times \Lambda^2$$

$$= \int_0^R D_s(N)N(r)2\pi r dr \Big/ \int_0^R N(r)2\pi r dr. \tag{58}$$

Allis and Rose give the quantity D_s/D_-. The integration can be performed graphically, to give $\langle D_s\rangle/D_-$ *vs* N_0, the electron density at the center of the cavity both radially and longitudinally. Figure 2 shows the quantities D_s/D_- and $\langle D_s\rangle/D_-$ *vs* $N_0\Lambda^2\mu_-/D_-$ computed in this manner.

The theoretical results for hydrogen are shown in Fig. 3, where contours of constant $N_0\Lambda^2$ are plotted on the $p/E_e - E_e\Lambda$ plane. The contour $N_0\Lambda^2=0$ represents breakdown, and the contour $N_0\Lambda^2=\infty$ represents ambipolar diffusion. At large values of $N_0\Lambda^2$ ($\gtrsim 10^8$) and intermediate to small values of p/E_e (<0.05) the calculations have not been made sufficiently exactly to warrant their inclusion in this figure. The difficulty may be traced to the approximation Eq. (37) made in solving the tail of the distribution function. The following parameters pertaining to hydrogen have been used in the calculation:

(a) $\nu_e = 4.85\times10^9 p$ sec^{-1} (p=mm Hg);

(b) $u_x = 8.9$ ev. This choice is explained by Ramien[10] on a quantum-mechanical basis.

(c) $u_i = 16.2$ ev.

(d) $h_x = h_i = 1.1\times10^{-2}$ (volts^{-1}). These quantities are adjusted so that the products $\nu_e h_x(u-u_x)$ and $\nu_e h_i(u-u_i)$ are in agreement with the published values of the probabilities of excitation and ionization.

(e) $\mu_+/\mu_- = 32$. The value $\mu_+ = 14.7$ cm^2/volt-sec at 760 mm Hg and 18°C given by Tyndall[11] is used. This value is constant over the dc range $0<E/p<20$ volts/cm-mm Hg, and should be independent of the microwave field. Allis and Rose show that the dc space charge E/p is less than the upper limit except in the immediate vicinity of the walls for a few combinations of high electron density and low pressure. Since $\mu_- = e/m\nu_e$ for constant ν_e we have at 18°C, in mks units, $\mu_+ = 1.12/p$, $\mu_- = 36/p$.

(f) $D_+\mu_-/D_-\mu_+ = 0$, an approximation justified by A-R.

Figure 4 shows the theoretical results of Fig. 3 plotted as a surface in $E_e\Lambda$, p/E_e, $N_0\Lambda^2$ space. The surface in the region of large $N_0\Lambda^2$ and small $E_e\Lambda$ has

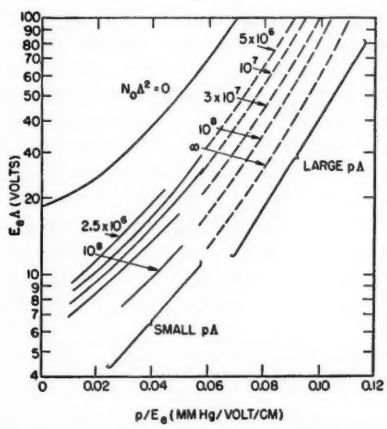


FIG. 3. Theoretical results for hydrogen where contours of constant $N_0\Lambda^2$ (in units of cm^{-1}) are plotted on the $p/E_e - E_e\Lambda$ plane.

[9] M. A. Herlin and S. C. Brown, Phys. Rev. **74**, 1650 (1948).

[10] H. Ramien, Z. Physik **70**, 353 (1931).

[11] A. M. Tyndall, *The Mobility of Positive Ions in Gases* (Cambridge University Press, Cambridge, 1938).

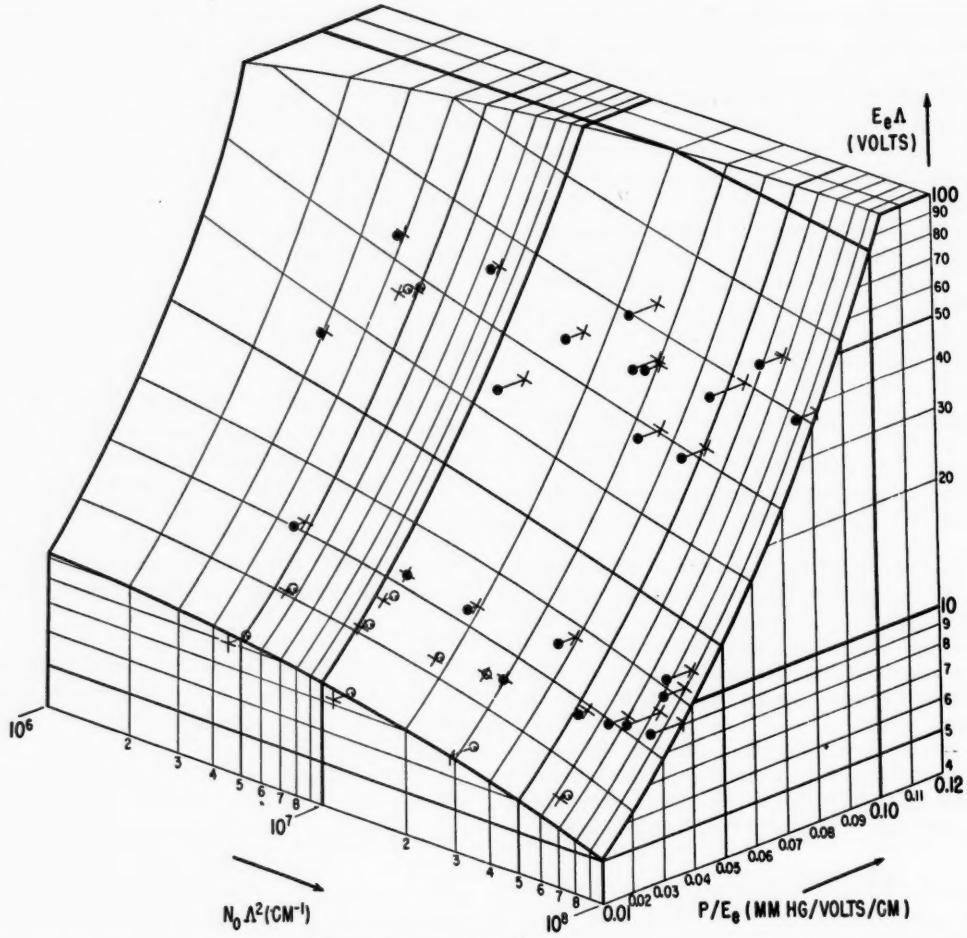


FIG. 4. A theoretical surface on which the experimental results are plotted. The experimental points are shown as dots, and their projections in the p/E_e direction into the theoretical plane as crosses.

been obtained by extrapolation of the theoretical calculations of Fig. 3. The experimental points are shown as dots; their projection in the p/E_e direction onto the theoretical plane, as crosses. The average of the experimental points is in good agreement with the theory (within 10 percent) both at large $E_e\Lambda$ and near $E_e\Lambda=10$. At the lowest values of $E_e\Lambda$ and p/E_e, the electronic mean free path is about equal to the diffusion length, and diffusion theory can no longer be applied. The discrepancy at intermediate values, amounting at worst to about 20 percent in p/E_e near the ambipolar limit, may be traced to the approximation made in solving the tail of the distribution function. In this region, near the ambipolar limit, the high pressure approximation begins to fail, and the asymptotic expansion Eq. (37) converges slowly. In each case, the error tends to raise p/E_e for a given $E_e\Lambda$. The experimental error is believed to be about six percent.

Ambipolar Diffusion in a Magnetic Field

ALBERT SIMON
Oak Ridge National Laboratory, Oak Ridge, Tennessee
(Received December 13, 1954)

Diffusion of ions in a plasma across a magnetic field is shown to be *not* ambipolar in character in most arc experiments. Owing to the highly anisotropic conductivity of the medium, the ions diffuse across the field at their own intrinsic rate. Space-charge neutralization is maintained by slight adjustments of the currents in the direction of the magnetic field lines. The discrepancy between theory and experiment noted by Bohm is thus resolved and no additional mechanisms, such as plasma oscillations, need be postulated.

IN the absence of a magnetic field, electrons would tend to diffuse out of a plasma much more readily than the ions. These unequal diffusion rates immediately produce large space charges and hence electric fields in the plasma so as to reduce the electron current, increase the ion current and maintain space charge neutralization. The resultant ambipolar diffusion coefficient D, which is the same for both electrons and ions, is of the order of magnitude[1]

$$D\cong 2D_+D_-/(D_++D_-), \tag{1}$$

where D_+ and D_- are the intrinsic ion and electron diffusion coefficients respectively. Since $D_-\gg D_+$, we have in the absence of a magnetic field

$$D\cong 2D_+, \tag{2}$$

and the diffusion proceeds approximately at the rate of the *slower* component.

The effect of a strong magnetic field is to reduce the coefficients of diffusion, $D_\perp$, across the magnetic field to the values

$$D_{+\perp}=\frac{D_+^0}{1+(\omega_+\tau_+)^2},\qquad D_{-\perp}=\frac{D_-^0}{1+(\omega_-\tau_-)^2}, \tag{3}$$

where the superscript zero denotes the coefficient for $H=0$, $\omega_\pm=eH/m_\pm c$, and $\tau_\pm=$mean free time between collisions for ions or electrons respectively. Diffusion in the direction of the magnetic field, $D_\parallel$, is unchanged from its field-free value. In most arc experiments $(\omega_-\tau_-)\gg(\omega_+\tau_+)\gg 1$ and as a result

$$D_{-\perp}\ll D_{+\perp}\ll D_\pm^0=D_{\pm\parallel}. \tag{4}$$

It might now be supposed, since the ions diffuse more rapidly across the magnetic field, that electric fields would arise which would reduce the ion current and which would result in a new ambipolar diffusion coefficient, as before, of the magnitude

$$D_\perp\cong 2D_{-\perp}. \tag{5}$$

Just such an analysis has been applied by Bohm[2] to experiments on the diffusion of ions in an arc plasma across a magnetic field. For the conditions of these arcs, such as occur in Calutron ion sources, $(\omega_-\tau_-\cong 10^5)$, Eq. (5) yielded an expected value of $D_\perp=20$ cm^2/sec. The observed values[3] were more like $D_\perp=3\times 10^3$ cm^2/sec. As a result of this large discrepancy, Bohm postulated that the principal mechanism of diffusion was by means of plasma oscillations. However, subsequent experimental attempts[4] to produce these oscillations have proven to be fruitless.

It is the purpose of this note to point out that diffusion across the magnetic field in all experiments of this type is *not* ambipolar in character. The discrepancy between theory and experiment noted by Bohm therefore does not exist and no additional mechanism, such as plasma oscillations, need be postulated.

The absence of ambipolar diffusion is due to the highly anisotropic conductivity in the medium. A section of an arc chamber is sketched in Fig. 1. Consider a fluctuation in which some positive ions begin streaming to the right, producing a region, 2, of net positive charge, and leaving a region, 1, of net negative charge to the left. An electric field immediately builds up to counteract this separation. However, this electric field has an enormously greater effect on the currents in the direction of the magnetic field lines because of the enormously greater $[(\omega\tau)^2\gg 1]$ conductivity of the medium in this direction compared to that across the field lines. As a result, space charge neutralization is maintained by slightly reducing the electron current from region 2 to the end wall and slightly increasing it from region 1 to the end wall. Of course, an electric field of the same order of magnitude is produced in the horizontal direction. This field is entirely too small to have any effect on the perpendicular diffusion because of the greatly reduced conductivity in this direction. One way of picturing the effect is to observe that the combined effect of the increased electron current to the wall from region 1, the presence of the end wall, and

[1] See M. A. Biondi and S. C. Brown, Phys. Rev. **75**, 1700 (1949).

[2] A. Guthrie and R. K. Wakerling, *The Characteristics of Electrical Discharges in Magnetic Fields* (McGraw-Hill Book Company, Inc., New York, 1949), National Nuclear Energy Series, Plutonium Project Record, Vol. 5, Div. 1, p. 197.

[3] See reference 2, p. 201, Eq. (12).

[4] D. H. Looney and S. C. Brown, Phys. Rev. **93**, 965 (1954).

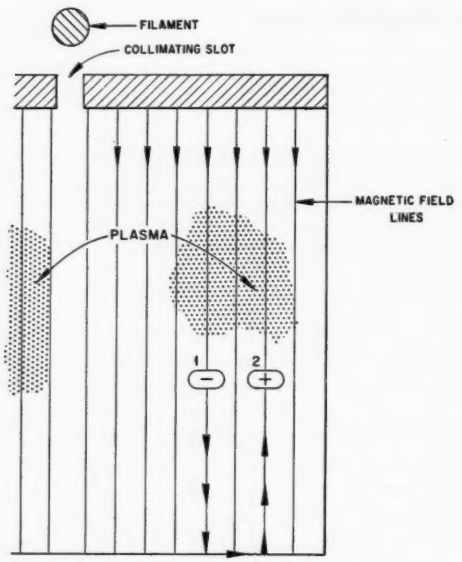


FIG. 1. Illustration of the "short-circuit" effect.

the decreased electron current from region 2 is equivalent to an electron "short circuit" which obviates the need for an ambipolar electric field in the perpendicular direction. As a result, electrons and ions do not diffuse at the same rate across the magnetic field, but rather with the coefficients given by Eq. (3). Space-charge neutralization is maintained by slight modification of the currents to the wall in the direction of the magnetic field lines.

Another way of looking at this effect is to note that in a medium with anisotropic conductivity it is no longer necessary that the currents flowing to the wall balance at the same equilibrium value in each direction. A positive excess to the side wall will be balanced by a negative excess to the bottom and top walls. As a consequence, ambipolar diffusion may be restored if the end walls are separately insulated and allowed to float.

A formal derivation of this effect may be obtained by revising the results of reference 2 so as to include the effect of the potential V, due to the deviation of the plasma from neutrality, upon the diffusion in the direction of the magnetic field lines. Equations (6) and (7) of reference 2 now become (in our notation)

$$D_{+\perp}\frac{d^2n_+}{dx^2}+\frac{D_{+\perp}e}{kT_+}\frac{d}{dx}\left(n_+\frac{dV}{dx}\right)+\frac{D_{+\parallel}e}{kT_+}\frac{d}{dz}\left(n_+\frac{dV}{dz}\right)=\frac{\beta n_+}{l}, \quad (6)$$

$$D_{-\perp}\frac{d^2n_-}{dx^2}-\frac{D_{-\perp}e}{kT_-}\frac{d}{dx}\left(n_-\frac{dV}{dx}\right)-\frac{D_{-\parallel}e}{kT_-}\frac{d}{dz}\left(n_-\frac{dV}{dz}\right)=\frac{\gamma n_-}{l}. \quad (7)$$

These equations are a statement of the conservation of charge for ions (n_+) and electrons (n_-) respectively. The x-direction is perpendicular to the magnetic field direction and the z-axis is parallel to this field. The quantities β and γ represent the mean velocity of diffusion of ions and electrons respectively along the field lines to the end walls of the arc chamber in the absence of the small perturbation due to the potential V.

The terms involving dV/dz were omitted in reference 2. However, these terms are many orders of magnitude larger than the terms involving dV/dx. To see this it is sufficient to recognize that $d[n_+dV/dx]/dx$ will be of the same order of magnitude as $d[n_+dV/dz]/dz$ while $(D_{+\parallel}/D_{+\perp})\cong(\omega_+\tau_+)^2\cong 10^5$ for the usual arc conditions. Hence the second term on the left side of Eqs. (6) and (7) may be neglected. Assuming space-charge neutralization, $n_+\cong n_-$, dV/dz can be eliminated between these equations, giving

$$(D_{+\perp}D_{-\parallel}T_++D_{-\perp}D_{+\parallel}T_-)\frac{d^2n_+}{dx^2}=\frac{n_+}{l}(\beta D_{-\parallel}T_++\gamma D_{+\parallel}T_-). \quad (8)$$

The solution of this differential equation which vanishes for large values of x is

$$n_+=A\exp\{-x/x_0\},$$

where

$$x_0=\left[\frac{l(D_{+\perp}D_{-\parallel}T_++D_{-\perp}D_{+\parallel}T_-)}{(\beta D_{-\parallel}T_++\gamma D_{+\parallel}T_-)}\right]^{\frac{1}{2}}. \quad (9)$$

Now $D_{+\perp}\gg D_{-\perp}$, $D_{-\parallel}\gg D_{+\parallel}$ and $\beta\cong\gamma$. Hence Eq. (9) may be rewritten as

$$x_0\cong[lD_{+\perp}/\beta]^{\frac{1}{2}}, \quad (10)$$

which differs from the result given in Eq. (11) of reference 2 in that the quantity $D_{+\perp}$ has replaced $(1+T_+/T_-)D_{-\perp}$.

The theoretically expected value of $D_{+\perp}$ as calculated from Eq. (3) for the conditions of the Bohm experiment is

$$D_{+\perp}\cong 3\times 10^3 \text{ cm}^2/\text{sec},$$

in agreement with the observed value.

Slope of the λ Curve of Liquid Helium*

K. R. Atkins
Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania
(Received December 22, 1954)

Landau's theory of liquid helium is developed to give an approximate value of the slope of the λ curve in terms of the quantities $(\rho/\Delta)\partial\Delta/\partial\rho$, $(\rho/p_0)\partial p_0/\partial\rho$, and $(\rho/\mu)\partial\mu/\partial\rho$. From the experimental data on the coefficient of expansion and the pressure variation of the velocities of first and second sound, the following numerical values are derived for these quantities:

$$\frac{\rho}{\Delta}\frac{\partial\Delta}{\partial\rho}=-0.57,\quad \frac{\rho}{p_0}\frac{\partial p_0}{\partial\rho}=+0.26,\quad \frac{\rho}{\mu}\frac{\partial\mu}{\partial\rho}=-1.8.$$

The negative slope of the λ curve can then be explained and its magnitude predicted with an accuracy of about 15 percent.

LANDAU'S latest theory of liquid helium[1] contains three arbitrary parameters, Δ, p_0, and μ, which occur in the expression relating the energy ϵ of a roton to its momentum p:

$$\epsilon=\Delta+(p-p_0)^2/2\mu. \tag{1}$$

The contribution of the rotons to the entropy can be shown to be

$$S_r=\frac{2k^{\frac{1}{2}}\mu^{\frac{1}{2}}p_0^2\Delta}{(2\pi)^{\frac{1}{2}}\rho T^{\frac{1}{2}}\hbar^3}\left(1+\frac{3kT}{2\Delta}\right)e^{-\Delta/kT}. \tag{2}$$

The contribution of the rotons to the fraction of normal component is

$$x_r=\rho_{nr}/\rho=\frac{2\mu^{\frac{1}{2}}p_0^4}{3(2\pi)^{\frac{1}{2}}\rho(kT)^{\frac{1}{2}}\hbar^3}e^{-\Delta/kT}. \tag{3}$$

Landau pointed out that the temperature of the λ point, T_λ, may be obtained approximately by putting $x_r=1$, or

$$\frac{2\mu^{\frac{1}{2}}p_0^4}{3(2\pi)^{\frac{1}{2}}\rho(kT_\lambda)^{\frac{1}{2}}\hbar^3}e^{-\Delta/kT_\lambda}=1. \tag{4}$$

Using the experimental data on the entropy, S, and the velocity of second sound, u_2, along the vapor pressure curve, Khalatnikov[2] deduced that $\Delta/k=8.9\pm0.2$°K, $p_0=(2.1\pm0.05)\times10^{-19}$ deg cm sec^{-1}, $\mu=(1.72\pm0.68)\times10^{-24}$ g. Inserting these values in Eq. (4) one obtains $T_\lambda=2.55$°K, as compared with the observed value of 2.19°K. Exact agreement is not to be expected, because the simple theory underlying Eq. (3) is invalid above 1.6°K where the mutual interactions of the rotons become important. However, it is clear that Eq. (4) is capable of giving the right order of magnitude for T_λ and it might be added that the rigorous mathematical treatment of λ transitions is now known to be so complicated that no simple theory could expect to make exact numerical predictions.

We might therefore hope to obtain an approximate value for the slope of the λ curve by differentiating Eq. (4) with respect to the density ρ. The result is

$$\frac{\rho}{T_\lambda}\frac{\partial T_\lambda}{\partial\rho}\left(\frac{\Delta}{kT_\lambda}-\frac{1}{2}\right)=\frac{\Delta}{kT_\lambda}\frac{\rho}{\Delta}\frac{\partial\Delta}{\partial\rho}-\frac{1}{2}\frac{\rho}{\mu}\frac{\partial\mu}{\partial\rho}-\frac{4\rho}{p_0}\frac{\partial p_0}{\partial\rho}+1. \tag{5}$$

From their measurements of the coefficient of expansion, Atkins and Edwards[3] deduced that

$$\frac{\rho}{\Delta}\frac{\partial\Delta}{\partial\rho}=-0.57, \tag{6}$$

$$\frac{1}{2}\frac{\rho}{\mu}\frac{\partial\mu}{\partial\rho}+\frac{2\rho}{p_0}\frac{\partial p_0}{\partial\rho}=-0.38. \tag{7}$$

In addition, it is necessary to know the value of $(\rho/p_0)\times\partial p_0/\partial\rho$. Khalatnikov[2] has quoted $(\rho/p_0)\partial p_0/\partial\rho\sim+\frac{1}{3}$, but gives no details of his calculations. He probably used a method similar to that given below, but his estimate must have been very rough since it was made before the recent measurements of entropy[4] and coefficient of expansion[3] and it is also not clear whether he made an adequate allowance for phonon effects.

To obtain $(\rho/p_0)\partial p_0/\partial\rho$, we start with the equation for the velocity of second sound:

$$u_2^2=\frac{\rho_s}{\rho_n}\frac{TS^2}{C}. \tag{8}$$

Differentiating with respect to the pressure,

$$\frac{1}{x(1-x)}\frac{\partial x}{\partial p}=-\frac{2}{u_2}\frac{\partial u_2}{\partial p}-\frac{2V\alpha_p}{S}+\frac{TV}{C}\left(\frac{\partial\alpha_p}{\partial T}+\alpha_p^2\right). \tag{9}$$

Hence, knowing the coefficient of expansion α_p as a function of temperature[3] and u_2 as a function of pressure,[5,6] $\partial x/\partial p$ can be derived. Here x is the total

* Work supported in part by the U. S. Atomic Energy Commission.

[1] L. Landau, J. Phys. (U.S.S.R.) **11**, 91 (1947).
[2] I. M. Khalatnikov, J. Exptl. Theoret. Phys. (U.S.S.R.) **23**, 8 (1952).
[3] K. R. Atkins and M. H. Edwards (to be published).
[4] G. R. Hercus and J. Wilks, Phil. Mag. **45**, 1163 (1954).
[5] V. P. Peshkov and K. M. Zinoveva, J. Exptl. Theoret. Phys. (U.S.S.R.) **18**, 438 (1948).
[6] R. D. Maurer and M. A. Herlin, Phys. Rev. **76**, 948 (1949).

fraction of normal component including both roton and phonon contributions. But the phonon contribution is

$$x_{ph}=TS_{ph}/c^2, \tag{10}$$

where c is the velocity of first sound. Hence

$$\frac{\partial x_{ph}}{\partial p}=-\frac{T}{c^2}\left(V\alpha_{ph}+2S_{ph}-\frac{1}{c}\frac{\partial c}{\partial p}\right), \tag{11}$$

and as S_{ph} and $(1/c)\partial c/\partial p$ are known,[7,8] $\partial x_{ph}/\partial p$ can be obtained and subtracted from $\partial x/\partial p$ to give $\partial x_r/\partial p$.

Differentiating (2) and (3) with respect to ρ, $(\rho/p_0)\times\partial p_0/\partial\rho$ may finally be determined from

$$2\frac{\rho}{p_0}\frac{\partial p_0}{\partial\rho}=\frac{1}{x_rK_T}\frac{\partial x_r}{\partial p}+\frac{\rho}{\Delta}\frac{\partial\Delta}{\partial\rho}\frac{1}{(1+3kT/2\Delta)}+\frac{V\alpha_r}{S_rK_T}, \tag{12}$$

where K_T, the isothermal compressibility, and α_r, the contribution of the rotons to the coefficient of expansion, are known.[3] Taking $(1/u_2)\partial u_2/\partial p$ from the measurements of Peshkov and Zinoveva,[5] $(\rho/p_0)\partial p_0/\partial\rho$ is found to be $+0.26$. However, this result is very sensitive to $(1/u_2)\partial u_2/\partial p$ and, if the measurements of Maurer and Herlin[6] are used, $(\rho/p_0)\partial p_0/\partial\rho$ varies from -1.12 at 1.0°K to -0.23 at 1.6°K. As it should, in principle, be independent of temperature, we have preferred to use the Russian measurements, but it would be desirable to make a further investigation of the variation of u_2 with pressure, placing particular emphasis on the lower pressures to obtain an accurate value for the initial slope. It is also important to note that all the parameters discussed here are sensitive to the values adopted for S and C. We have used the recent values of Hercus and Wilks[4] which are about 10 percent higher than previous values.

If we accept $(\rho/p_0)\partial p_0/\partial\rho=+0.26$, Eqs. (6) and (7) lead to $(\rho/\mu)\partial\mu/\partial\rho=-1.8$. It appears that the effective mass μ varies much more rapidly with ρ than either Δ or p_0.

Equation (5) may now be evaluated to yield $(\rho/T_\lambda)\times\partial T_\lambda/\partial\rho=-0.42$. The experimental value is -0.37. In view of the approximation involved in neglecting roton interactions, the agreement is very satisfactory. It seems that both the negative coefficient of expansion and the negative slope of the λ curve are caused primarily by the negative value of $(\rho/\Delta)\partial\Delta/\partial\rho$.

[7] Kramers, Wasscher, and Gorter, Physica **18**, 329 (1952).
[8] K. R. Atkins and R. A. Stasior, Can. J. Phys. **31**, 1156 (1953).

PHYSICAL REVIEW VOLUME 98, NUMBER 2 APRIL 15, 1955

Phenomenological Theory of Townsend Breakdown in Dielectrics

P. L. AUER
General Electric Research Laboratory, Schenectady, New York
(Received December 7, 1954)

The Townsend model of dielectric breakdown is discussed. The transient equations describing the behavior of the electron and positive ion current densities are solved rigorously with boundary conditions appropriate to a simplified Townsend model employing photoelectron production at the cathode as the only secondary mechanism. The correct boundary condition is found to lead to a set of integro-difference equations. Solution of the integro-difference equations allows rigorous specification of the criterion for breakdown. This criterion is found to become equivalent to the familiar Townsend criterion in the asymptotic limit of large time intervals. The results of our development may be applied to the calculation of formative time lags and to the estimation of space-charge distortion effects.

1. INTRODUCTION

WHEN a dielectric material is subjected to an electrical stress of sufficient magnitude it is converted into a conductor. Breakdown occurs when the conductivity of the material becomes sufficient to maintain a current limited only by the external circuit. In his pioneering work on electrical discharges Townsend found that in most instances the conversion of an insulator to conductor is caused by the ionization of the original dielectric material.

The ionization of the dielectric can be initiated by any of several well-known external means, or by the presence of stray electrons in the system. The electrons created by the initial ionization burst become accelerated in the applied electric field and in turn produce additional ionization. Such an avalanche of ionization can, however, produce no breakdown as long as the field remains static and essentially homogeneous since the electrons created in the ionization avalanche are removed from the system by the applied field. In order to achieve a self-sustaining discharge, it becomes necessary to introduce a secondary, regenerative source of electron production; although it is conceivable that in some rare instances breakdown can be achieved purely by some highly effective field distortion mechanism instead of a secondary electron production process.

Various schemes have been proposed to explain the secondary electron production process active in the breakdown of dielectrics by static fields. We shall confine our attentions to what we shall term Townsend

breakdown, and we define this phenomenon as the breakdown of dielectric material in the presence of a *homogeneous static* electric field where the only secondary electron producing process active takes place at the cathode surface. In excluding from our analysis the possibility of ionization by positive ions or photons in the bulk of the dielectric, we do not necessarily imply that such processes may not be effective in secondary electron production. However, we are of the opinion that the mechanism we term Townsend breakdown plays an important role in many breakdown studies and a complete analysis of this model may prove to be valuable aid in elucidating the actual mechanisms governing breakdown.

In what follows, we shall present a detailed analysis of the transient growth of the ionization current in a Townsend breakdown and subsequently derive the criterion for breakdown. For the sake of convenience in mathematical operation an especially simplified model has been chosen for the initial analysis. We shall assume that the only active secondary electron production process present is due to the action of photons at the cathode surface and that these photons, created in the primary ionization avalanche, travel through the dielectric material without absorption and with infinite speed. In later sections of this presentation, we shall discuss modifications necessary to include the action of positive ions at the cathode surface and the consequences of photon absorption and finite speeds of travel.

2. DIFFERENTIAL EQUATION OF BREAKDOWN

a. Equation of Continuity

We consider a quantity of dielectric placed between two infinite plane parallel electrodes at δ cm apart. At a given initial time, a voltage of constant magnitude is impressed across the electrodes producing a homogeneous static field in the space occupied by the dielectric. We imagine that at the instant the electric stress is applied to the dielectric there arises an initial current due to the action of the field on either stray electrons present in the system or electrons introduced from an external source. For the sake of convenience in mathematical operation, we restrict the initial current to the origin of our coordinate system, which is conveniently placed at the cathode. As we shall show later, no intrinsic difficulty is met with in considering the initial current to have some arbitrary distribution in space. The number of tedious mathematical operations is greatly increased, however, if we depart from the simple localized picture of the initial current which we shall presently develop in some detail.

We shall find it possible to obtain a complete description of the electron current in terms of the following quantities: α, defined as the relative increase in current density per cm in the direction of the field, β, defined here as the secondary ionization coefficient and used to denote the efficiency of photoelectron production at the cathode; and v, which is the electron drift velocity and will be treated as a vector quantity. These three quantities are functions of the nature of the particular dielectric under consideration. In addition, α is a sensitive function of the electric field strength while β may be influenced by the nature of the cathode material. For the system under discussion, the drift velocity is assumed to be constant in magnitude and direction for any given field strength. We choose a Cartesian frame with the x axis placed in the direction of the field. The cathode is placed at the origin and the anode is placed at $x=\delta$ cm. Denoting the time variable by t, the initial time is chosen such that $t=0$ when the field is first applied and the initial current makes its appearance at the cathode.

We let $i=i(x,t)$ represent the electron current density. Then for the system under discussion the equation of continuity takes the particularly simple form

$$\frac{1}{v}\frac{\partial i}{\partial t}+\frac{\partial i}{\partial x}=\alpha i, \tag{2.1}$$

where v is the drift velocity of the electron and α is the primary ionization coefficient. The solution of the above differential equation must satisfy a boundary condition incorporating the secondary electron generating process. The boundary condition for the system under discussion is given by the relation

$$i(0,t)=i(0,0)+\beta\int_0^{\delta} i(x,t)dx, \tag{2.2}$$

where β is the secondary ionization coefficient for photoelectron production at the cathode.

The system of equations presented by (2.1) and (2.2) is the one appropriate to discussing Townsend type breakdown in dielectrics. It is essentially the same system that was treated by previous investigators in the field.[1–4] The differences between our model and those employed by the authors cited above will be elaborated on later. For the most part the remainder of this paper will be concerned with presenting a detailed solution of the above system of equations.

b. Steady-State Solution

Before proceeding to the main portion of our investigation, it will be instructive to consider solutions to the system of Eqs. (2.1) and (2.2) at steady state. These solutions are well known and described in most texts and articles on dielectric breakdown. The results for steady state will be made use of in subsequent discussion.

At steady state the time variation disappears and

[1] W. Bartholomeyczyck, Z. Physik **116**, 235 (1940).
[2] H. L. von Gugelburg, Helv. Phys. Acta **20**, 250, 307 (1947).
[3] P. M. Davidson, Brit. J. Appl. Phys. **4**, 173 (1953).
[4] H. W. Bandel, Phys. Rev. **95**, 1117 (1954).

the equation of continuity for the electron current density reduces to

$$di/dx=\alpha i, \quad i=i(x); \tag{2.3}$$

along with the boundary condition

$$i(0)=i_0+\beta\int_0^\delta i(x)dx. \tag{2.4}$$

In Eq. (2.4) the quantity i_0 must be interpreted as the hypothetical steady-state value of the cathode electron current in the absence of a secondary mechanism.

The solution to the system of Eqs. (2.3) and (2.4) is given by

$$i(x)=\frac{i_0e^{\alpha x}}{1-(\beta/\alpha)[e^{\alpha\delta}-1]}, \tag{2.5}$$

as may be verified by direct substitution. Townsend was the first to derive the expression given in Eq. (2.5) on the basis of physical arguments. Recognizing that physically this expression predicts a catastrophe whenever the denominator becomes zero, Townsend interpreted the onset of the catastrophe as the criterion for dielectric breakdown. Consequently, it has become customary practice to define breakdown in terms of the relation

$$(\beta/\alpha)[e^{\alpha\delta}-1]=1. \tag{2.6}$$

It must be recognized that there is no reason to believe steady-state conditions can ever prevail at the point of breakdown and the process of analysis leading to Eq. (2.6) as the criterion for breakdown must be deemed unsatisfactory. The proper criterion for breakdown should be obtained from the solution of Eq. (2.1) instead of (2.3). We shall now proceed to this task.

c. Transient System

The solution of Eq. (2.1) can be written down readily in the form

$$i(x,t)=f(t-x/v)e^{\alpha x}, \tag{2.7}$$

where $f(z)$ is an arbitrary function of its argument in so far as the differential equation is concerned but is required to satisfy Eq. (2.2) in order to represent a solution to our problem. On setting x equal to zero in (2.7) it is seen that $f(t)$ is simply the cathode electron current density at time t.

Substituting the expression (2.7) into Eq. (2.2) we obtain

$$f(t)=f(0)+\beta\int_0^\delta f\left(t-\frac{x}{v}\right)e^{\alpha x}dx; \tag{2.8}$$

and on rearranging terms under the integral sign we obtain the integral equation

$$f(t)=f(0)+\beta v\int_{t-\delta/v}^{t} e^{\alpha v(t-s)}f(s)ds. \tag{2.9}$$

The solution of the above equation represents the central problem in characterizing Townsend breakdown and is the major topic of our investigation.

3. INTEGRAL EQUATION OF TOWNSEND BREAKDOWN

a. System of Integral Equations

The form of Eq. (2.9) suggests that it represents a set of difference equations. This may be demonstrated in the following manner. The transit time of an electron, t_e, from one electrode to the other is simply $\delta/|v|$. The time axis may then be divided in units of electron transit times. We define a set of functions according to the relation

$$\begin{gathered} f(t)\equiv\phi_n(y), \quad n=1, 2, \cdots, \\ (n-1)t_e\leqslant t\leqslant nt_e, \\ 0\leqslant y\leqslant 1, \\ t_e=\delta/|v|. \end{gathered} \tag{3.1}$$

Setting the initial time equal to zero, Eq. (3.1) yields the matching conditions:

$$\begin{aligned} \phi_0(y)&=0, & n&=0 \\ \phi_n(0)&=\phi_{n-1}(1), & n&>0. \end{aligned} \tag{3.2}$$

With the aid of the above definitions, Eq. (2.9) may be written in the form

$$\phi_n(y)=f(0)+p\int_y^1 e^{\sigma(y-z)}\phi_{n-1}(z)dz+\lambda\int_0^y e^{\sigma(y-z)}\phi_n(z)dz, \tag{3.3}$$

$$p\equiv\lambda e^\sigma, \quad \lambda\equiv\beta\delta, \quad \sigma\equiv\alpha\delta.$$

The quantities, p, λ, σ have been introduced for the sake of convenience in notation.

From Eq. (3.3), we obtain the particular relations:

$$\begin{aligned} \phi_n(1)&=f(0)+p\int_0^1 e^{-\sigma z}\phi_n(z)dz, \\ \phi_n(0)&=f(0)+p\int_0^1 e^{-\sigma z}\phi_{n-1}(z)dz, \end{aligned} \tag{3.4}$$

and consequently we check that the integral equation (3.3) satisfies the matching conditions of (3.2). On differentiating Eq. (3.3) with respect to y we obtain

$$\phi_n'(y)=q\phi_n(y)-p\phi_{n-1}(y)-\sigma f(0), \quad q\equiv\lambda+\sigma, \tag{3.5}$$

where the quantity q has been introduced for the sake of convenience in notation and the prime denotes differentiation with respect to the argument. An additional matching condition may be obtained by combining (3.2) with (3.5) to yield

$$\phi_n'(0)=\phi_{n-1}'(1), \quad n>2. \tag{3.6}$$

Thus, the $\phi_n(y)$ represent a function of time which after the second transit time interval is not only continuous but also smooth.

It is interesting to note that the system of equations described by (3.3) bears some resemblance to systems encountered in criticality calculations of nuclear reactors. The resemblance is not too surprising if we recognize that the Townsend model under consideration is a sort of chain reaction itself.

b. Solution for the Initial Time Span

Equation (3.3) takes a particularly simple form for the initial electron transit-time span. Since $\phi_0(y)$ doesn't exist,

$$\phi_1(y)=f(0)+\lambda\int_0^y e^{\sigma(y-z)}\phi_1(z)dz. \tag{3.7}$$

This is an elementary inhomogeneous Volterra integral equation of the second kind. It is instructive to note that with the exception of pathological cases the homogeneous Volterra equation possesses no nonzero solutions. The above is not an exception to this rule and consequently Eq. (3.7) can have no solution other than zero unless $f(0)$, which is the cathode electron current density at zero time, has a value different from zero. We stress this point merely to refute the contention of previous authors that the Townsend model for breakdown can be described by a spontaneous creation equation where the electron current is finite for times greater than the initial time but is zero initially.

Equation (3.7) is readily solved by any of several familiar methods to yield

$$\phi_1(y)=f_0\{1+(\lambda/q)[e^{qy}-1]\}.^5 \tag{3.8}$$

c. Solution for General Time Spans

The solution of the set of equations given in (3.3) may be accomplished by several methods. In Appendix 1 of this presentation we describe a direct solution of the set of integro-difference equations by means of a simple operational technique. Immediately below we shall sketch briefly a solution of the analogous differential-difference equations given by (3.5).

We define two functions in the following manner:

$$\zeta_n(y)=\phi_n(y)-\left(\frac{\sigma}{q-p}\right)f_0, \quad q\neq p,$$
$$G(s,y)=\sum_{j=1}\zeta_j(y)s^{j-1}, \tag{3.9}$$

where it is seen that $G(s,y)$ is a generating function and the power series is assumed to converge for small enough values of s. On substituting $\zeta_n(y)$ for $\phi_n(y)$ into Eqs. (3.5) and (3.2), multiplying by appropriate powers of s, and adding, we obtain

$$\frac{\partial G(s,y)}{\partial y}=(q-ps)G(s,y)+\frac{p\sigma}{q-p}f_0,$$
$$G(s,0)=sG(s,1)+\left(1-\frac{\sigma}{q-p}\right)f_0, \quad q\neq p. \tag{3.10}$$

The solution of the above equation is given by

$$G(s,y)=-\left(\frac{p\sigma}{q-p}\right)\frac{f_0}{q-ps}$$
$$+(1-se^{q-ps})^{-1}\left(1-\frac{\sigma}{q-ps}\right)f_0e^{(q-ps)y}. \tag{3.11}$$

Expansion of the above expression for $G(s,y)$ as a power series in s and association of the appropriate coefficients in the expansion provides the required solution for the $\phi_n(y)$ of Eq. (3.5) with the restriction $q\neq p$.

If we set y equal to zero in Eq. (3.11), the following expression is obtained:

$$\frac{\phi_n(0)}{\phi_1(0)}=\frac{\sigma}{q}\left[\frac{1-(p/q)^n}{1-(p/q)}\right]+\frac{\lambda}{q}A_{n-1}$$
$$-\frac{\sigma}{q}\sum_{j=0}^{n-2}A_j(p/q)^{n-j-1}, \tag{3.12}$$

where the A_j are the coefficients of s^j in the expansion of the generating function $[1-se^{(q-ps)}]^{-1}$ and are given by the expressions

$$A_0=1, \quad A_n=\sum_{j=0}^{n-1}(-1)^j\frac{(n-j)^j}{j!}p^je^{(n-j)q}. \tag{3.13}$$

The expression for $\phi_n(y)$ may be written in the form

$$\frac{\phi_n(y)}{\phi_1(0)}=\frac{\sigma}{q}\left[\frac{1-(p/q)^n}{1-(p/q)}\right]+e^{qy}\sum_{j=0}^{n-1}(-1)^j\frac{(py)^j}{j!}\epsilon_{n-j},$$
$$\epsilon_n=(\lambda/q)A_{n-1}-(\sigma/q)\sum_{j=0}^{n-1}A_j(p/q)^{n-j-1}. \tag{3.14}$$

The results of Eqs. (3.12) and (3.14) may be applied to the case $p=q$ if we make use of the relation

$$\lim_{x\to 1}\frac{1-x^n}{1-x}=n. \tag{3.15}$$

Equations (3.12) to (3.14) represent the required solution to the original integral equation of (2.9). Unfortunately the sums involving the coefficients A_j

[5] Reference 4 quotes the results of some calculations made by W. Kunkel. For the model under discussion here, that is, where the only secondary mechanism active is photoelectron production at the cathode, Kunkel's results would be in agreement with our Eq. (3.8) providing his role of positive ion and photon mechanisms were reversed. The time range of validity quoted in reference 4 would then be incorrect, however. Actually, Eq. (10) of reference 4 appears to predict that for our model, steady-state conditions are obtained after a single electron transit time has occurred.

have no convenient representation in closed form and the exact calculation of the $\phi_n(y)$ becomes a tedious tabulation process. We may, however, make use of the following properties obtained in usual practice:

$$pe^{-q}\equiv\lambda e^{-\lambda},\quad \lambda\ll 1. \tag{3.16}$$

After some calculation we obtain from (3.12) and (3.13)

$$\frac{\phi_n(0)}{\phi_1(0)}=\frac{\lambda}{q}\left(\frac{p}{q}\right)^n+\frac{\sigma}{q}\left[\frac{1-(p/q)^n}{1-(p/q)}\right]+p^nO(\lambda^n). \tag{3.17}$$

For any practical purpose the last term on the right of Eq. (3.17) may be neglected with negligible loss in accuracy. Figure 1 presents the behavior of the cathode electron current density with time according to the relation given by (3.17). The curves were obtained by smoothly joining together points obtained from (3.17).

4. CRITERION FOR BREAKDOWN AND THE APPROACH TO STEADY STATE

a. Physical Criterion

It can be shown easily on the basis of Eq. (3.3) that for real positive values of the coefficients α and β the cathode electron current, $f(t)$, must be a monotonic increasing function of time. In analogy with Townsend it may be asked whether there are values which may be assigned to α and β which will provide for arbitrarily large currents in the course of time regardless of the initial current value. In order to ascertain whether such values exist for α and β it is necessary to examine the behavior of Eq. (3.14) in the asymptotic limit of large "n."

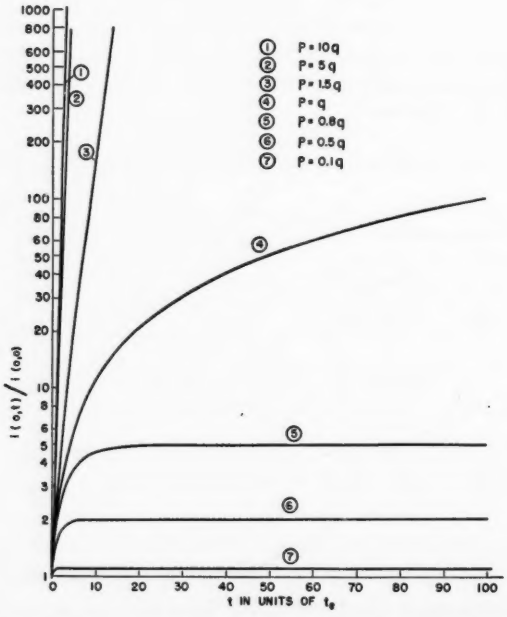


FIG. 1. Cathode electron current density as a function of time, according to the relation given by Eq. (3.17).

There is an alternative and simpler method available for deriving the breakdown criterion. We state that a system is capable of undergoing breakdown only if the current is unable to approach a finite asymptotic value, which may be associated with its steady-state value. This requires that the first time derivative of the current can never become a decreasing function of time. However, it is sufficient to require that the derivative be a constant, and we shall use this as the concept for breakdown criterion.

In particular, the constancy of the first time derivative permits us to require

$$\phi_{n+1}'(0)=\phi_n'(0) \tag{4.1}$$

for all values of "n." Equation (4.1) will now be developed to provide the criterion for breakdown.

b. Mathematical Criterion for Breakdown

From Eq. (3.12) we obtain the relation

$$\phi_{n-1}(0)=\left(\frac{p}{q}\right)\phi_n(0)+\phi_1(0)\left\{\frac{\lambda}{q}A_n-\left(\frac{p}{q}\right)A_{n-1}+\frac{\sigma}{q}\right\}. \tag{4.2}$$

Inserting (4.2) into (3.5) provides

$$\phi_{n+1}'(0)-\phi_n'(0)=\lambda\phi_1(0)\{A_n-(1+e^q)A_{n-1}+e^qA_{n-2}\}=0. \tag{4.3}$$

The above equation constitutes a set of finite difference equations for the A_n; the necessary boundary conditions may be obtained from (3.13):

$$A_0=1;\quad A_1=e^q. \tag{4.4}$$

The solution to Eq. (4.3) subject to (4.4) is readily obtained by the usual methods to yield

$$A_n=1+(e^q-1)\left[\frac{1-e^{n\sigma}}{1-e^{\sigma}}\right]. \tag{4.5}$$

It can be readily shown that the coefficients A_n must satisfy the relation

$$\sum_{j=0}^{n}(-1)^ja_jA_{n-j}=0,\qquad a_0=1,\quad a_n=e^qp^{(n-1)}/(n-1)!, \tag{4.6}$$

since it will be recognized that the a_j are simply the expansion coefficients of $(-s)^j$ in the generating function $[1-se^{(q-p\sigma)}]$. Substitution of (4.5) into (4.6) yields

the criterion for breakdown in the form

$$\left(\frac{e^q-e^\sigma}{1-e^\sigma}\right)\left[1-e^q\sum_{k=0}^{n}(-1)^k\frac{p^k}{k!}\right]$$
$$+e^{n\sigma}\left(\frac{e^q-1}{e^\sigma-1}\right)\left[1-e^\lambda\sum_{k=0}^{n}(-1)^k\frac{\lambda^k}{k!}\right]=0, \quad (4.7)$$
$$q=\lambda+\sigma, \quad p=\lambda e^\sigma.$$

In Appendix B we show that the second term of Eq. (4.7) can be made to approach zero as n approaches infinity. Consequently, the criterion for breakdown becomes asymptotically

$$1-e^{(q-p)}=0, \quad (4.8)$$

which requires that

$$p=q, \quad \beta e^{\alpha\delta}=\alpha+\beta. \quad (4.9)$$

Equation (4.9) is the familiar Townsend criterion cited previously in (2.6). It would have been possible to obtain this criterion by inspection of Eq. (3.17) with the sacrifice of some rigor.

c. Approach to Steady State

It has been established in the previous section that for values of (p/q) less than unity the system can approach a limiting steady-state condition. Since

$$\lim_{n\to\infty}\left[\frac{1-x^n}{1-x}\right]=\frac{1}{1-x}, \quad x<1, \quad (4.10)$$

Eq. (3.14) provides

$$\lim_{n\to\infty}\phi_n(y)=\binom{\sigma}{q}\frac{\phi_n(0)}{1-p/q}$$
$$=\frac{f_0}{1-(\beta/\alpha)(e^{\alpha\delta}-1)}, \quad (4.11)$$
$$p/q<1.$$

Comparison of the above result with Eq. (2.5) verifies that (4.11) predicts the correct asymptotic value for the cathode electron current density in the event the system is below the breakdown threshold.

The value of the cathode electron current for the case $p=q$ can be evaluated readily using Eqs. (3.17) and (3.15). We obtain

$$\phi_n(0)=\left(\frac{n\sigma}{q}+\frac{\lambda}{q}\right)\phi_1(0)+O(n\lambda); \quad p=q. \quad (4.12)$$

5. ELECTRON AND POSITIVE ION CURRENT DISTRIBUTION IN SPACE

a. Electron Current

In Sec. 3. a. we have shown that the electron current density at the cathode is described for given times by the $\phi_n(y)$ according to the relation of (3.1). Making use of Eq. (2.7) we find that the spatial distribution of the electron current density at a given time may be written as

$$i(x,t)/i(0,t)=e^{\alpha x}\phi_n(y-x/\delta)/\phi_n(y), \quad (5.1)$$
$$\phi_n(-z)=\phi_{n-1}(1-z), \quad 0<z<1.$$

In particular the electron current density at the anode becomes

$$i(\delta,t)=\phi_{n-1}(y)e^\sigma, \quad (n-1)t_e\leqslant t\leqslant nt_e. \quad (5.2)$$

The above relation confirms the fact that no electron current can develop at the anode until a single electron transit time has elapsed.

b. Positive Ion Current

The positive ion current density is denoted $j(x,t)$ and is found to satisfy a relation analogous to (2.1) in the form

$$\frac{1}{w}\frac{\partial j}{\partial t}+\frac{\partial j}{\partial x}=-\alpha i, \quad (5.3)$$

where w is the drift velocity of the positive ion and all other symbols have been defined previously. The solution to the above equation has to satisfy the boundary condition

$$j(\delta,t)=0. \quad (5.4)$$

Solution to Eqs. (5.3) and (5.4) may be constructed in the form

$$i(x,t)=0, \quad (t-|x/v|)<0,$$
$$j(x,t)=\alpha\int_0^\delta i(s;t-x/w+s/w)ds, \quad (t-|x/v|)\geqslant 0, \quad (5.5)$$
$$1/u=1/v-1/w,$$

where $i(x,t)$ is given by Eq. (2.7).

In particular the positive ion current density at the cathode is given according to Eqs. (5.5) and (2.7) by the relation

$$j(0,t)=\alpha u\int_{t-\delta/u}^{t}f(s)e^{\alpha u(t-s)}ds, \quad (5.6)$$

where $f(t)$ is given by Eq. (2.9).

We shall have no occasion in this presentation to calculate specifically the positive ion current density. However, we wish to point out that the results developed in previous sections permits us to calculate both the electron and positive ion current density distribution in space for arbitrary times. Having made these calculations we can then proceed to find the resulting electric field, $E(x,t)$, from the relation

$$\partial E/\partial t=j(x,t)+i(x,t)+C(t), \quad (5.7)$$

and $C(t)$ may be obtained from the requirement

$$V_0=-\int_0^\delta E(x,t)dx, \tag{5.8}$$

where V_0 is the value of the externally impressed voltage, and we have used mks units in the above equations. Equation (5.7) provides a description of the field under the assumption that field distortions due to space charge are negligible. The consistency of this argument may be tested by solving for the field according to (5.7).

6. DISCUSSION

a. Generalized Townsend Model

In the previous sections we have developed in detail the description of the electron and positive ion current in a Townsend discharge system where the only active secondary mechanism present was the photoelectron production at the cathode. Several other forms of secondary mechanisms have been considered instrumental in Townsend discharge. For present purposes we need only consider secondary mechanisms occurring at the cathode. There are then two important alternatives to consider, photoelectron production and electron production due to positive ion impact at the cathode.

We may generalize and consider that both photons and positive ions are effective in producing electrons at the cathode. The efficiency of electron production due to photons will be denoted by β as before, while the efficiency due to the positive ions will be denoted by the quantity ω/α. We are still required to solve Eq. (2.1) but the boundary condition of (2.2) is now replaced by

$$i(0,t)=i(0,0)+\frac{\omega}{\alpha}j(0,t)+\beta\int_0^\delta i(x,t)dx. \tag{6.1}$$

Making use of Eqs. (2.7) and (5.6), the above may be written

$$f(t)=f(0)+\omega u\int_{t-\delta/u}^{t} e^{\alpha u(t-s)}f(s)ds + \beta v\int_{t-\delta/v}^{t} e^{\alpha v(t-s)}f(s)ds. \tag{6.2}$$

With $\omega=0$, Eq. (6.2) becomes equal to (2.9). With $\beta=0$, it becomes analogous to (2.9) and the solution proceeds in the same manner, the role of v being replaced with u. When the quantities ωu and βv have comparable values, Eq. (6.2) must be solved as stands. The equation may be converted into an integro-difference equation analogous to (3.3) providing the electron transit time $t_e(t_e\equiv\delta/|v|)$ is an integral multiple of the positive ion transit time $t_p(t_p\equiv\delta/|\omega|)$. The solution of (6.2) is then carried out in the same manner as the development presented in Sec. 3.

Additional difficulties are encountered if the simplified model chosen here for discussion is made more realistic. For example, we have assumed the initial electron current is restricted to the cathode. If at the initial time there is a finite distribution of electrons in space, we are required to construct relations analogous to Eq. (2.2) for each localized source of initial current and obtain the net current by superposition. In the event the dielectric material under consideration has an appreciable photon absorption coefficient, Eq. (2.2) must be amended accordingly. The effect of absorption causes no undue difficulty in treatment only as long as the coefficient is a well-defined function of space and time. The fact that photons travel with finite speeds instead of infinite, as we have tacitly assumed, may also be accounted for by amending Eq. (2.2); this correction, however, is negligible in most cases of physical interest.

b. Formative Time Lag

The application of the results given in Sec. 3 to the calculation of formative time lags is obvious. It is only necessary to find the time intervals required by the relation (3.17) to reach some arbitrary value adjudged equal to the value of the electron current density at the cathode during breakdown. In order to compare our results with experimental data it would be highly desirable to estimate accurately the magnitude of the breakdown current at either the cathode or anode.

APPENDIX A.

Consider the integral operator

$$\mathbf{I}=\int_0^y ds e^{\sigma(y-s)}. \tag{A-1}$$

Equation (3.7) of the text may be written in terms of $\mathbf{I}$ as

$$(1-\lambda\mathbf{I})\phi_1(y)=f_0; \tag{A-2}$$

and the solution of this equation provides the relation

$$(1-\lambda\mathbf{I})^{-1}=(1+\lambda\mathbf{J}),\quad \mathbf{J}=\int_0^y ds e^{q(y-s)}. \tag{A-3}$$

After adding and subtracting the term $p\mathbf{I}\phi_{n-1}(y)$ on the right of (3.3) we obtain

$$(1-\lambda\mathbf{I})\phi_n(y)=f_0-p\mathbf{I}\phi_{n-1}(y) + e^{\sigma y}[\phi_n(0)-\phi_1(0)], \tag{A-4}$$

where use has been made of (3.4). Application of the operational algebra of (A-3) to (A-4) yields the difference equation

$$\phi_n(y)-\phi_n(0)e^{qy}=\phi_1(y)-\phi_1(0)e^{qy}-p\mathbf{J}\phi_{n-1}(y); \tag{A-5}$$

and the solution of (A-5) is given by

$$\phi_n(y)=\sum_{k=0}^{n-1}(-1)^k p^k \mathbf{J}^k\left\{\phi_1(0)\frac{\sigma}{q}[1-e^{qy}]+\phi_{n-k}(0)e^{qy}\right\}, \quad \text{(A-6)}$$

where use has been made of Eq. (3.8).

The following relations involving the operator $\mathbf{J}$ are now required:

$$\mathbf{J}^m(ae^{qy})=\frac{ay^m}{m!}e^{qy},$$

$$\mathbf{J}^m(a)=(-1)^{m-1}aq^{-m}e^{qy}\sum_0^{m-1}(-1)^k\frac{(qy)^k}{k!}+(-1)^m aq^{-m}, \quad \text{(A-7)}$$

a=constant.

Substituting the above results into (A-6) and combining terms yields

$$\phi_n(y)=\frac{\sigma}{q}\phi_1(0)\sum_{k=0}^{n-1}(p/q)^k+\sum_{k=0}^{n-1}(-1)^k\frac{(py)^k}{k!}\phi_{n-k}(0)e^{qy}-\frac{\sigma}{q}\phi_1(0)\sum_{k=0}^{n-1}\sum_{h=0}^{k}(-1)^h(p/q)^k\frac{(qy)^h}{h!}e^{qy}. \quad \text{(A-8)}$$

The first sum or the right side of the above equation is a geometric series. We introduce the following notation:

$$R_n(\lambda,\sigma)\equiv\frac{\sigma}{q}\sum_{k=0}^{n-1}(p/q)^k=\frac{\sigma}{q}\left[\frac{1-(p/q)^n}{1-p/q}\right], \quad p\neq q, \quad \text{(A-9)}$$

and note the following useful property of the polynominal $R_n(\lambda,\sigma)$:

$$R_{n-m}=(p/q)^{-m}[R_n-R_m]. \quad \text{(A-10)}$$

The order of summation in the last term of (A-8) is now inverted and the results of (A-9, 10) are applied to obtain the final form

$$\phi_n(y)-\phi_1(0)R_n(\lambda,\sigma)=e^{qy}\sum_{k=0}^{n-1}(-1)^k\frac{(py)^k}{k!}\times[\phi_{n-k}(0)-\phi_1(0)R_{n-k}(\lambda,\sigma)]. \quad \text{(A-11)}$$

A set of relations may be obtained for the $\phi_j(0)$ on setting $y=1$ in expression (A-11). Using the notation

$$\phi_1(0)\Psi_n=\phi_n(0)-\phi_1(0)R_n(\lambda,\sigma), \quad \text{(A-12)}$$

we obtain the set of equations

$$\sum_{k=0}^{n-1}(-1)^k a_k\Psi_{n-k}=\theta_n,$$

$$\begin{cases}a_0=1, & a_k=e^q p^{(k-1)}/(k-1)!,\\ \theta_1=\lambda/q, & \theta_j=-(\sigma/q)(p/q)^{j-1}.\end{cases} \quad \text{(A-13)}$$

The set of linear equations given in (A-13) may be inverted by standard matrix methods. We note, however, that the coefficients a_k are the coefficients of $(-1)^k s^k$ in the expansion of the generating function $F(s)=1-se^{(q-ps)}$. Consequently, we may obtain the elements of the triangular matrix inverse to the matrix of a_k by collecting the coefficients of s^k in the expansion of $[1/F(s)]$. Proceeding in this manner, we obtain

$$\Psi_n=\sum_{j=0}^{n-1}A_j\theta_{n-j},$$

$$A_0=1, \quad \text{(A-14)}$$

$$A_n=e^{nq}\sum_{j=0}^{n-1}(-1)^j\frac{(n-j)^j}{j!}(pe^{-q})^j.$$

Equations (A-11) and (A-14) correspond to Eqs. (3.14) and (3.12) and (3.13) of the text, respectively.

APPENDIX B.

We are indebted to Dr. H. Poritsky for suggesting the method of proof described below. It is required to show that

$$\lim_{n\to\infty}e^{nu}\left\{1-e^s\sum_0^n(-1)^j s^j/j!\right\}=0, \quad u>s. \quad \text{(B-1)}$$

We write, for complex s and t,

$$\sum_0^n(-1)^j s^j/j!=\frac{1}{2\pi i}\oint dt e^{-st}\sum_1^{n+1}t^{-j}=\frac{1}{2\pi i}\oint dt e^{-st}\left[\frac{1-t^{-(n+1)}}{t-1}\right], \quad \text{(B-2)}$$

where the contour is taken to include the origin and unity on the positive real axis. According to the theory of residues,

$$e^{-s}=\frac{1}{2\pi i}\oint\frac{e^{-st}}{t-1}dt, \quad \text{(B-3)}$$

the contour being extended to infinity. Combining results, we find

$$\lim_{n\to\infty}e^{nu}\left\{1-e^s\sum_0^n(-1)^j s^j/j!\right\}=\lim_{n\to\infty}\frac{e^{-u}}{2\pi i}\oint\left(\frac{e^u}{t}\right)^{n+1}\frac{e^{-s(t-1)}}{(t-1)}dt. \quad \text{(B-4)}$$

Since u is finite and e^{-st} is analytic for positive t at least up to infinity, the quantity on the right of (B-4) may be made to approach zero by choosing t large enough. This completes the required proof.

Helium II Film Transport. I. The Role of Substrate*

BERNARD SMITH† AND HENRY A. BOORSE‡
Pupin Physics Laboratories, Columbia University, New York, New York
(Received January 11, 1955)

Liquid helium II film transport from aluminum, silver, copper, nickel, nickel-silver, stainless steel, Pyrex, quartz, and Lucite beakers has been measured using a variation of a previously described electrical technique. The degree of reproducibility and background variation observed in separate experiments on the same or on similarly prepared specimens of the same material was studied in detail. Rates on machined metals were found to be considerably lower than would be expected from many previous reports. In fact, rates measured over all the aforementioned materials were so similar in magnitude and reproducibility that the background variations involved prevent the resolution of any systematic dependence on substrate material which might exist. This is consistent with the size of the helium-substrate interaction as calculated by Schiff for copper, silver, and glass.

INTRODUCTION

ACCORDING to the two-fluid model of liquid helium II, only superfluid atoms are capable of flowing through narrow channels whose width is of the order of 10^{-4} cm or less. Since the thickness of the creeping film is considerably smaller than channels which may conveniently be constructed, flow through the film appears to be the nearest approximation to ideal superfluidity and should present very favorable conditions for the study of superflow in liquid helium II.

In spite of the widespread attention that the phenomenon of film flow has received in laboratories throughout the world, study of the literature reveals not only lack of agreement on quantitative theoretical predictions, but also a more serious lack of agreement on the quantitative aspects of many experimental results. Owing to the difficulties presented by such measurements, many of the reported results depend on a rather meager amount of data. Moreover, a proper evaluation of conflicting reports is frequently hampered by insufficient discussion of experimental details due to the brevity imposed on journal publications. Thus the progress of the theory tends to be retarded by the lack of agreement upon the precise observations which must be explained, while the experimental work suffers from the concomitant lack of sound theoretical guidance.

The present studies of the bulk transport rates of the helium II film are concerned with the clarification of the answers to a number of basic and closely related questions. The first of these may be stated as follows: How do the measured transport rates vary with the nature of the substrate? The answer to this question has varied considerably since the problem was first considered in 1938 by Daunt and Mendelssohn, who concluded that "the rate of transfer is a characteristic quantity which only depends on the thermal state of liquid helium II."[1] Thus they remarked that "although a higher rate of transfer per unit surface was observed on drawn copper wires, the rate on polished copper was found to be exactly the same as on glass. We attribute the higher rate on drawn wires to the influence of surface inhomogeneities, and conclude that the rate of transfer is not influenced by the underlying material."[1] From these experiments it was also concluded that the high "transfer over a given surface can be reduced by making the surface smooth."[2]

After the war the subject was re-examined and, on the basis of new experiments with glass, platinum, and nickel beakers, it was reported that the absolute values of the rate of transfer "on metals were found to be much higher than on glass,"[3] or more specifically, "the conclusion therefore seems inescapable that the transfer is actually higher on a surface of platinum or nickel than on glass."[3] The situation was further complicated by the very large range of transport rates reported by various observers for metals, in comparison with the much smaller range of values reported for glass (see Table I).[4–18] Indeed, in an article[10] accompanying the

* Assisted by the Office of Naval Research and the Linde Air Products Company.
† Present address: Bell Telephone Laboratories, 463 West Street, New York 14, New York.
‡ Barnard College, Columbia University, New York 27, New York.

[1] J. G. Daunt and K. Mendelssohn, Nature **142**, 475 (1938).
[2] J. G. Daunt and K. Mendelssohn, Proc. Roy. Soc. (London) **A170**, 423 (1939).
[3] K. Mendelssohn and G. K. White, Proc. Phys. Soc. (London) **A63**, 1328 (1950).
[4] J. G. Daunt and K. Mendelssohn, Proc. Roy. Soc. (London) **A170**, 439 (1939).
[5] Webber, Fairbank, and Lane, Phys. Rev. **76**, 609 (1949).
[6] H. A. Fairbank and C. T. Lane, Phys. Rev. **76**, 1209 (1949).
[7] B. V. Rollin and F. Simon, Physica **6**, 219 (1939) with the added explanation contained in reference 6.
[8] K. R. Atkins, Proc. Roy. Soc. (London) **A203**, 241 (1950).
[9] G. J. van den Berg, Report of Proceedings of Oxford Conference on Low Temperature Physics, (August, 1951), p. 69.
[10] J. G. Daunt and K. Mendelssohn, Proc. Phys. Soc. (London) **A63**, 1305 (1950).
[11] B. N. Eselson and B. G. Lazarev, J. Exptl. Theoret. Phys. (U.S.S.R.) **23**, 552 (1952). We are grateful to Dr. B. S. Chandrasekhar for supplying us, in April, 1954, with a translation of this paper prepared at the University of Illinois.
[12] G. S. Picus, Phys. Rev. **94**, 1459 (1954).
[13] J. G. Dash and H. A. Boorse, Phys. Rev. **82**, 851 (1951).
[14] W. C. Knudsen and J. R. Dillinger, Phys. Rev. **91**, 489 (1953).
[15] E. Ambler and N. Kurti, Report of Proceedings of Oxford Conference on Low Temperature Physics, (Aug. 1951), p. 70.
[16] L. C. Jackson and D. G. Henshaw, Phil. Mag. **41**, 1081 (1950).
[17] B. S. Chandrasekhar and K. Mendelssohn, Proc. Phys. Soc. (London) **A65**, 226 (1952).
[18] B. S. Chandrasekhar, Phys. Rev. **86**, 414 (1952).

TABLE I. Summary of previously reported transport rates. Since these rates vary slowly with temperature between 1°K and 1.5°K, the tabulated values may be compared even for slightly different temperatures in this range. All rates quoted in the table are usually assumed to have been obtained for uncontaminated specimens.

Material	T°K	Rate in cm³/cm-sec ×10⁴	Observational technique	Reference	Remarks
Glass	1–1.5	7.5	Visual	2, 4[b]	
Glass	1.4	9.1	Visual	5[b]	
Glass	1.4	12.3	Vaporization	6[b]	
Glass	1.6	~8	Vaporization	7	
Glass	1.47	7.5–8.7[a]	Visual	8[b]	
Glass	?≤1.5	7.5–12[a]	Visual	8[b]	
Glass	?≤1.5	8.6–14.3[a]	Visual	8[b]	
Glass	1.5	7.1	Visual	3[b]	
Glass	1.5	8.6	Visual	3[b]	Baked at 130°C
Glass	?≤1.5	8	Gravimetric	9	
Glass	?≤1.5	9	Gravimetric	9	
Glass	?≤1.5	11	Gravimetric	9	Baked at 300°C
Glass	1.25	7.5	Electronic	13	
Glass	1.5	9	Visual	12[b]	
Glass	1.5	~9–12[a]	Visual	11[b]	
Copper	1–1.5	~7.5	Visual	2, 4, 10	
Copper	1.3	14.8	Electronic	13[b]	Machined
Copper	1.3	23	Electronic	13[b]	Burnished
Copper	1.3	51.2	Electronic	13[b]	Etched
Copper	1.3	26	Electronic	13[b]	Etched and oxidized
Copper	1.3	19	Visual	14	
German silver	1.6	~8	Vaporization	7	Tubing
German silver	1.5	~8.5	Vaporization	15[b]	Tubing
Stainless steel	1.3	10.4	Electronic	13[b]	Machined
Stainless steel	1.3	31	Visual	14	
Stainless steel	1.3	16.9	Visual	16[b]	Barium stearate on outside surface
Stainless steel	1.3	8.3	Visual	17[b]	Polished
Stainless steel	1.3	26	Visual	17[b]	Polish damaged by bringing to "red heat"
Nickel	1.5	22.7	Visual	3[b]	Polished, baked
Iron	1.3	10.5	Electronic	13[b]	Machined
Platinum	1.5	16.9	Visual	3[b]	Drawn tubing, acid cleaned
Platinum	1.5	20.8	Visual	3[b]	Baked tubing
Platinum	1.5	21.0	Visual	3[b]	Polished, baked tubing
Lead	1.5	16	Electronic	13[b]	Oxidized
Lucite	1.3	14	Electronic	13[b]	Rouge polished, uncorrected for contraction
Lucite	1.3	11.5	Visual	18[b]	Rouge polished
Perspex	1.3	10.3	Visual	18[b]	Rouge polished
Perspex	1.3	15	Visual	18[b]	Liquid polished

[a] Height dependence noted in these data.
[b] These references contain rate *vs* temperature or height dependence data for this particular material (in this connection see papers III and IV of the present series).

one in which rates "on metals were found to be much *higher* than on glass,"[3] one finds that the reported properties of the thermomechanical effect were based on measurements and calculations in which the transport rate over copper wires was *identical* with that over glass surfaces. In the light of this background some observers have been led to the conclusion that "glass appears to be the only surface over which simple and reproducible phenomena of film flow have been observed."[19]

One widespread interpretation of these conflicting reports has been based on the supposition that, in contrast to glass, ordinary metal surfaces (e.g., machined, drawn, etc.) are characterized by a high degree of surface roughness[1,2] which acts either to present a microperimeter many times larger than the measured macroperimeter,[14,20] or else to produce surface cracks which enhance the flow of bulk liquid.[3,17] According to this view, "there may be only one transfer rate (that exhibited by the glass surfaces) and any higher rates would be due to the microstructure of the substrate."[17] The finding of Chandrasekhar and Mendelssohn that highly polished stainless steel yielded a rate "practically the same as that for a baked-out glass surface"[17] as observed by Mendelssohn and White,[3] is frequently cited in support of this interpretation.[20] Aside from the experimental conditions to which we shall have occasion to refer in paper II of this series, these results (in the absence of a pre-polishing control measurement) do not seem sufficiently different from previous measurements on unpolished, machined iron and stainless steel[13] (or indeed from other results which have been reported for metals from time to time),[1,2,10,15] to justify definitive conclusions concerning the effect of surface polish. Furthermore, just as low rates have been reported for unpolished metals,[7,10,13,15] so also have observations on polished metals on occasion yielded high rates.[3,13]

[19] D. F. Brewer and K. Mendelssohn, Phil. Mag. **44**, 340 (1953).
[20] J. G. Daunt and R. S. Smith, Revs. Modern Phys. **26**, 172 (1954).

Furthermore, there are situations[14,16,21,22] (in addition to the one already cited[10]) in which it is desirable to determine a significant physical quantity from a calculation involving a separate knowledge of the film transport rate. For example, the film thickness, transport rate, and flow velocity are thought to be so related,[4,8,14,16,20–23] that knowledge of two of these quantities may be used to infer the third. However, the present state of knowledge is so unsatisfactory that in a recent publication such calculations are omitted because "until the present controversy as to the interpretation of the observed creep rates is finally settled it is perhaps inopportune to publish values of the calculated film velocities."[21]

The present paper is therefore concerned with an investigation of the variation of transport rate with substrate material. Since, as has already been noted, proper evaluation of such a study requires a knowledge of the influence of surface roughness on transport rates, the second paper in this series (designated as II) will describe control experiments on the role of surface finish. In combination, these studies will in turn be seen to require consideration of the roles of film height and temperature; consequently the influence of the latter parameters will be considered in papers III and IV.

EXPERIMENTAL TECHNIQUES AND APPARATUS[24]

I. Basis of the Electrical Method

A high-frequency electrical technique, using basic circuits which have already been described,[13,25] was adapted for use in these experiments by redesign of of capacitor transport vessels.[26] The dielectric properties of liquid helium are used to detect the helium liquid level in a cylindrical capacitor. Changes in the amount of liquid present in these capacitors are accompanied by frequency changes which ultimately manifest themselves as variations in the voltage which is fed to a recording potentiometer.[25]

II. Modified Capacitor Design

The results for metals were obtained with combined transport vessel-capacitors wherein each metal beaker also comprises one electrode of the capacitor. This eliminates the previous need to fabricate an extra cylindrical vessel of precisely known dimensions. Furthermore, this arrangement decreases the volume of liquid to be emptied from each beaker, thereby permitting measurements to be made more rapidly. However, in the case of dielectric specimens, where this technique is inapplicable, separate capacitor depth gauges[25] were used.

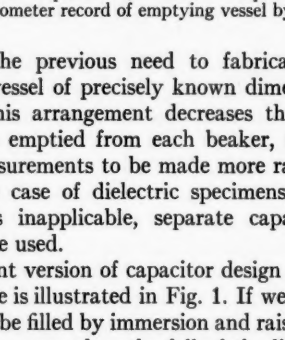


FIG. 1. Schematic representation of basic capacitor design and potentiometer record of emptying vessel by creep.

The present version of capacitor design and measuring technique is illustrated in Fig. 1. If we suppose the capacitor to be filled by immersion and raised to permit emptying by creep, then the fall of the liquid level in the region where the inner member (core) is composed of teflon (1) does not cause any change in capacitance. However, the arrival of the annular meniscus at the top of the metallic central section of the core is heralded by the onset of changes in capacitance which persist until the liquid level reaches the Teflon base (2), in which region changes in capacitance cease. The frequency-proportional voltage fed to the Brown Recording Potentiometer reflects regions of constant capacitance (and frequency) as traces parallel to the time axis, and regions of variable capacitance as traces inclined to the time axis, the angle being determined by the potentiometer chart speed, voltage sensitivity, and the transport rate. Thus the potentiometer chart, driven at a known, constant speed, serves as a clock providing a direct measure and permanent record of the time Δt consumed in emptying the known annular volume associated with the height Δh in the capacitor. These quantities, when correlated with the appropriate limiting perimeter lead directly to a determination of the rate. Thus, if in Fig. 1, $2b$ represents the inner diameter of the capacitor shell and $2a$ the diameter of the metallic section of the core, the transport rate R, given by the volume transferred in unit time per unit width of path, may be expressed as

$$R=[\bar{r}(\Delta r)(\Delta h)]/b(\Delta t), \tag{1}$$

where $\bar{r}=(a+b)/2$ and $(\Delta r)=b-a$.

The present capacitor design has the following advantages over that previously reported: (a) Although frequency is used as the basic observable, no actual measurement of frequency is required since only the temporal location of discontinuities in the voltage trace is employed quantitatively. The previously required calibration of the "frequency dependence of the potentiometer circuits"[13] is completely eliminated, and the transport rate data so obtained are independent of

[21] L. C. Jackson and D. G. Henshaw, Phil. Mag. **44**, 14 (1953).
[22] J. G. Dash, Phys. Rev. **94**, 1091 (1954).
[23] W. C. Knudsen and J. R. Dillinger, Phys. Rev. **95**, 279 (1954).
[24] This discussion is intended to be sufficiently general to serve to eliminate the need for further description of apparatus in subsequent papers in this series.
[25] H. A. Boorse and J. G. Dash, Phys. Rev. **79**, 734 (1950).
[26] B. Smith and H. A. Boorse, Proceedings of Schenectady Cryogenics Conference (General Electric Research Laboratory, The Knolls, Schenectady, 1952), p. 34; Phys. Rev. **90**, 156 (1953); Phys. Rev. **92**, 505 (1953). These brief progress reports were issued at stages in the research when the information they contained appeared to be of immediate interest. The present paper, inasmuch as it represents the first comprehensive report on this work, supersedes these. Slight numerical discrepancies between the present and previous reports stem from a careful recalculation of the preliminary results to remove sources of error.

the frequency characteristics and stability of the electronic circuits. (b) In the previous work each new depth gauge required calibration based on simultaneous frequency readings and cathetometer observations to permit the conversion of the observed time rate of change of frequency to the time rate of change of the position of the liquid level.[13] The present capacitor design is "self-calibrating" since the arrival of the liquid level at various heights in the capacitor is detected automatically during the course of a transport measurement.

III. Beakers and Radiation Shield

Figure 2 illustrates the important features of the actual capacitor and radiation shield assembly used in most of the measurements to be described. Although some of the results reported in this paper were obtained using earlier versions of this design which differed somewhat from Fig. 2 in scale and in the omission of some Teflon bushings, the design shown in Fig. 2 has proven so convenient that most of the data to be reported in the present series of papers were secured using this type of capacitor.[24]

The schematic, single "live-section" core of Fig. 1 has been replaced by one containing four sections (labelled 1–4), capped by Teflon bushing T–1 and supported by Teflon bushing T–5, both of which are secured to threaded projections of the core C. The core is machined in one piece, with the narrow connections between "live" sections encircled by Teflon bushings T–2, T–3, and T–4, having longitudinally split walls permitting them to be spread open and "snapped" into the positions shown. It may readily be seen that the potentiometer trace in emptying such a "four section beaker" consists of a fourfold repetition of the basic pattern of Fig. 1. Application of Eq. (1) to each metal section in turn leads to an average transport rate corresponding to the appropriate level in the beaker. Teflon bushing T–5, which fits snugly in the beaker serves not only to provide the final discontinuity in the potentiometer trace, but also provides insulation and alignment for the core in the shell S. Space for liquid helium in the vicinity of T–5 is provided by longitudinal grooves cut in the Teflon at convenient intervals on the circumference of the bushing. The transport vessel S was always machined in one piece in order to avoid spurious effects due to superleaks at the base.

Electrical connection is made to the core at a and to the shell at b, using double Teflon covered copper wire (#36 B&S) soldered with Wood's metal to the projections machined for this purpose.

The nut N serves to fasten the capacitor to the radiation shield. In the diagram, the shield and the capacitor are in electrical contact, whereas in some applications a special Teflon insulator replaced N and fastened the beaker to the shield while providing electrical isolation.

Typical capacitors had shells of $\frac{1}{2}$-in. o.d. with

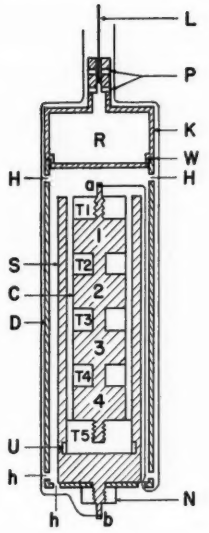


FIG. 2. Cross section of typical transport vessel-capacitor assembly in radiation shield (not to scale).

0.050-in. wall thickness. Using the notation of Eq. (1), other typical nominal sizes were $\Delta r = 0.005$ in. and $\Delta h = 0.350$ in., with beakers of 2.4-in. over-all inner depth. Further specifications of dimensions will be found in the data pertaining to each particular beaker.

Transport over dielectric surfaces was measured by using capacitor depth gauges which differ from the capacitor shown in Fig. 2 only with respect to the provision of holes in the base to insure equality of the helium levels in the depth gauge and dielectric beaker.[13] The extension of Eq. (1) to the case of a capacitor gauge within a beaker is straightforward.

Radiation shields of copper and brass were employed on different occasions. The shield (containing the capacitor) was suspended by a nylon thread L, to which the two electrical leads from a and b were loosely fastened at intervals of a few inches to provide a parallel wire line leading out of the cryostat (Fig. 3). The cap K, soldered to the shield body at W (Fig. 2) with Wood's metal, was provided with a reservoir R of liquid helium to reduce any possible temperature gradients.

IV. Determination of Annular Volume in Capacitors

Accurate measurement of the geometrical quantities appearing in Eq. (1) presents a problem only with respect to the value of b which, in contributing both to $\bar{r}$ and Δr, exerts its greatest influence on the calculated transport rate via Δr, since b and a are of the same order of magnitude. Bore measurement at various heights in the shell was accomplished with a special Brown and Sharpe "Intrimik" (Internal Tri-point Micrometer) which is read while in position; the tool's three jaws are retractable so that the micrometer may be inserted and removed from the hole without touching the walls between measurements. Thus the actual

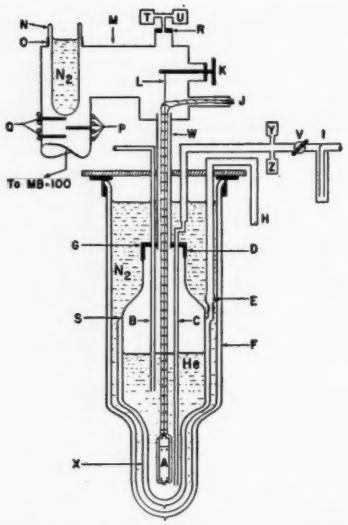


FIG. 3. Helium cryostat: Schematic representation of significant components (not to scale) showing: *A*: radiation shield assembly shown in Fig. 2; *B*: liquid helium filling tube; *C*: vapor pressure tube; *D*: copper-glass seal; *E*: Kovar-glass seal and bellows leading to Dewar vacuum line; *F*: liquid nitrogen Dewar; *G*: solder joint to copper cap; *H*: helium Dewar vacuum line; *I*: cold trap; *J*: multiple Kovar electrical seal for capacitor leads; *K*: vacuum tight metal winch; *L*: nylon thread; *M*: structure coupling helium bath to MB-100 diffusion pump backed by Kinney pump; *N*: Pyrex cold trap; *O*: Kovar-glass seal; *P*: cooling water circulation coils; *Q*: brass baffles; *R*: detachable cap for inserting specimens; *S*: ring seal of helium Dewar; *T*: thermocouple vacuum gauge; *U*: ionization gauge; *V*: stopcock; *W*: thin walled stainless steel tube; *X*: helium Dewar tail; *Y*: thermocouple vacuum gauge; *Z*: ionization gauge.

variation of bore with height is measurable to 0.0001 in. without damage to the walls, even in the case of a slightly tapered hole which enlarges with distance from the rim. The possibility of such measurement appears to be a distinct advantage of the geometry employed herein over that employed in conventional visual determinations of transport from capillaries about 1 mm in diameter.

Wherever necessary, appropriate corrections for contraction on cooling to liquid helium temperatures have been applied on the basis of published data for metals[27,28] and dielectrics.[29–31] The need for such corrections was minimized by making each capacitor core and shell from the same material.

The problem of accurate determination of the annular volume has been emphasized because of the importance of eliminating uncertainty due to geometric factors in interpreting any variations of transport rate with height in the beaker which might be observed (see paper III). Due to such uncertainty in the volume occupied by the Teflon bushings in Fig. 2, quantitative calculations of transport rate were restricted to the metallic sections labelled 1–4.

V. Helium Cryostat and Temperature Measurement

Figure 3 shows the radiation shield assembly *A* in the tail of the helium Dewar *X*.

Temperatures were determined from measured vapor pressures according to the "agreed scale of 1948."[32] The vapor pressure tube *C* is connected, via a copper-glass seal to the Pyrex stopcock *V*, which in turn leads to a liquid nitrogen cooled trap *I* before communicating with various pressure regulating and measuring devices. These include a mercury contact monostat,[13,33] permitting maintenance of constant vapor pressure, a 12 mm bore absolute mercury manometer, a calibrated Dubrovin vacuum gauge covering the range 0–20 mm Hg with ninefold magnification, and a special McLeod gauge reading pressures from 4.4 mm Hg to 2 microns. These gauges were supplemented by the thermocouple gauge *Y* (RCA–1946) which was calibrated against the McLeod gauge during runs.

VI. General Pre-Run Procedure

Inasmuch as contamination free operation is highly desirable in these measurements (see papers II and III) these procedures must be described in some detail.

Capacitor elements were washed in pure acetone and blown dry in a stream of pure helium gas before assembly and installation in the cryostat. As a final step, glass and quartz beakers were heated to dispel moisture. According to McBain, "heating glass in air at atmospheric pressure always lessens the quantity of gas evolved upon subsequent vacuum heat treatment, even though the glass may be under normal atmospheric conditions for an interval of several months between the two processes."[34]

Installation was invariably followed by immediate evacuation of the helium Dewar chamber using a Kinney pump which backs the MB–100 diffusion pump specified in Fig. 3. If this operation did not betray the presence of an obvious leak, the chamber was successively charged with pure helium gas and re-pumped to pressures of a few microns, this cycle being repeated five or more times. The pressures in this range were observed with thermocouple gauges located at *T* and *Y*.

This dilution of foreign gases was followed by pumping to pressures of about 2×10^{-5} mm Hg as read with an ionization gauge (type *VG–1A*) at *U*. A similar gauge *Z* is also provided near the stopcock *V*. All newly made joints were carefully tested at these pressures, even if the rate of evacuation appeared normal.

[27] F. C. Nix and D. MacNair, Phys. Rev. **60**, 597 (1941); Phys. Rev. **61**, 74 (1942); Rev. Sci. Instr. **12**, 66 (1941).
[28] H. Adenstadt, Ann. Physik **26**, 69 (1936).
[29] W. F. Giauque *et al.*, Rev. Sci. Instr. **23**, 169 (1952).
[30] H. L. Laquer and E. L. Head, U. S. Atomic Energy Commission Report 2161, 1952 (unpublished).
[31] F. A. Molby, J. Opt. Soc. Am. **39** 600 (1949).
[32] H. van Dijk and D. Shoenberg, Nature **164**, 151 (1949).
[33] J. J. Beaver and D. J. Beaver, J. Ind. Eng. Chem. **15**, 359 (1923).
[34] J. W. McBain, *The Sorption of Gases and Vapours by Solids* (Routledge, London, 1932).

TABLE II. Identification of beakers. Specimen symbols are assigned for convenient shorthand reference in the text. The measurement sequence may be inferred from the numbers assigned to runs; dates associated with these runs are indicated in subsequent graphical representation of the data.

Beaker material	Specimen symbol	Surface condition	Run numbers referred to in text
Aluminum, Alcoa alloy 11*S*, 5.5% Cu, 0.5% Pb, 0.5% Bi	Al-I (*B*)	Bored	38, 39, 48
	Al-II (*B*)	Bored	55
Copper, oxygen free high conductivity	Cu	Bored	59, 66
Glass:			
Pyrex, Kovar base	G-I	Untreated	9, 10
Pyrex, precision bore	G-II	Untreated	30, 31
Pyrex, precision bore	G-III	Untreated	56
Lucite	...	see reference 13	21
Nickel, Type "A," 99% pure	Ni-II (*R*)	Reamed	53
Nickel silver, 46.5% Cu, 40.75% Zn, 2.75% Pb, 10% Ni	Ni-Ag-I (*B*)	Bored	33–35, 37
	Ni-Ag-II (*B*)	Bored	45, 46, 49
Quartz			
Fused, transparent	Q-I	Untreated	23, 27, 32
Fused, transparent capillary	Q-II-V	Untreated	26, 27
Fused, translucent	Q-III	Untreated	24, 25
Silver, fine (99.9% pure)	Ag	Bored	58
Stainless steel, Type 303, 18% Cr, 8% Ni, 2% Mn, 1% Si	S. St. (*B*)	Bored	41–44, 47

The experimental chamber was then filled with pure helium gas, in which condition the apparatus remained overnight. The apparatus was then outgassed at about 2×10^{-5} mm Hg for periods which varied from about eight hours to several days before each helium run. No systematic correlation of transport rates with the duration of the outgassing period was observed.

The stopcock *V* was ordinarily closed and was opened only when the trap *I* was immersed in liquid nitrogen to prevent mercury vapor from entering the experimental chamber. With *V* closed, the pressure gauges beyond the trap *I* could be evacuated, leak tested, and recharged with helium gas using a separate pumping system and gas line.

VII. General Run Procedure and Performance of Apparatus

A single charge of liquid helium (approximately ½ liter at the boiling point) consistently proved sufficient for about 15 hours of transport rate measurement. Typical values of heat influx to the helium bath, inferred from flowmeter measurements throughout each run, varied from about 15 cal/hr with the bath at the level indicated in Fig. 3, to about 7 cal/hr when the level was in the Dewar tail. Transport rates were measured in successive cycles of decreasing and increasing temperatures to determine the possible effects of variations in bath level or to disclose progressive contamination during the run. No such effects were ever observed.

During transport measurements the lower portion of the radiation shield was always immersed in the helium bath to provide a chamber with each extremity in thermal contact with liquid helium at the same temperature. In addition to the transport rate data, further evidence of the efficacy of such shielding was provided by a number of special exploratory runs using carbon resistance thermometers,[35] with and without such helium capped shields. As expected, these tests showed that the shielding remained effective even when contact with the helium bath was temporarily broken.

Capacitors were designed to yield transport times of 100 sec or more; these were then measurable to within one second from the Brown Recording Potentiometer chart which was usually driven at a speed of two inches per minute. The chart speed was checked continuously against a Standard Electric Timer (model S–10), which could be read to the nearest 0.1 second. As in previous work,[13] interruption of the rf signal had no effect on the observed transport rates, thus reconfirming the absence of dielectric losses.

EXPERIMENTAL RESULTS

When this investigation was begun in 1950, it appeared from the most recent data then available[3,13,16,36,37] that transport rates varied considerably with substrate and/or surface finish.[38] Since the reported variations did not seem to follow any recognizable pattern, the present survey was undertaken in an effort to discern any systematic variation of transport rates with substrate which might exist. This would of course constitute an essential prelude to a study of the influence of surface finish (see paper II).

It was recognized that the wide range of previously reported rates (Table I) would make the results of such experiments very difficult to interpret unless the bounds of reproducibility characterizing such measurements were established. Many experiments were therefore devoted to the determination of the extent of the "unavoidable background" variation which could be

[35] Brown, Zemansky, and Boorse, Phys. Rev. **84**, 1050 (1951).
[36] K. Mendelssohn and G. K. White, Nature **166**, 27 (1950).
[37] H. A. Boorse and J. G. Dash, Phys. Rev. **79**, 1008 (1950).
[38] C. T. Lane, Ann. Revs. Nuclear Sci. **1**, 413 (1952).

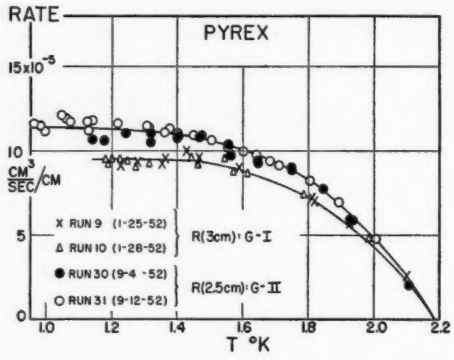


FIG. 4. Transport rates over Pyrex. Specimen G-I: averaged over distance from 1.9 cm to 4.1 cm from rim (lower curve); Specimen G-II: averaged over distance between 1.5 cm and 3.6 cm from rim (upper curve).

expected if one merely removed a beaker from the cryostat, dismantled the capacitor and radiation shield, and then re-subjected them to standard installation procedures before the next run. Unless otherwise specified, this procedure was always followed between consecutive runs with the same specimen. Due attention was also directed to the effect of merely "ageing" beakers outside the cryostat between runs. As a final precaution, transport was also measured from different specimens of the same material.

Beaker specimens are identified for future reference in Table II. The standard machining operation consisted of careful turning of the outer surfaces and boring (*B*) of the interiors, except in the case of specimen Ni-II where difficulty in boring to the proper tolerances necessitated the substitution of the reaming (*R*) process. No polishes or abrasives were used to modify the microfinish imparted by the initial machining operations. The copper, silver, aluminum, and nickel silver beakers were turned and bored without the use of any cutting fluid except for the sparing application of pure acetone. All metal beaker materials had to meet the obvious requirements of ready availability, satisfactory machinability, and reasonable corrosion resistance. Among those possessing these qualifications, further selection was based on various combinations of the desire for comparison with previous measurements on the same material, for comparison with theoretical estimates, and for materials which would also be suitable for use in later control experiments on the role of surface finish (see paper II).

Glass

Pyrex Specimen G-I

The lower curve in Fig. 4 presents data obtained in two runs with a Pyrex beaker whose base consisted of a Kovar-glass seal into which a close fitting Kovar plug was soft-soldered. This plug served to establish electrical connection to the shell of a nickel depth gauge. The beaker was left undisturbed in a helium atmosphere in the cryostat, except for repumping between the two runs, thereby providing information on reproducibility under these conditions. The data represent rates averaged over the distance from 1.9 cm to 4.1 cm from the beaker rim. It was evident that the Kovar seal base began to leak in the course of later work, and so no further measurements were performed with beakers which relied on solder joints or glass to metal seals for leak tightness.

Pyrex Specimen G-II

Since none of the data in the literature had described transport over glass of precision bore, measurements were undertaken with a beaker constructed from commercial precision bore tubing. The base of this beaker was carefully fused and annealed; the lower region, distorted in the fabrication of the beaker, did not overlap the sensitive region of the nickel depth gauge. The results, shown in the upper curve of Fig. 4 represent

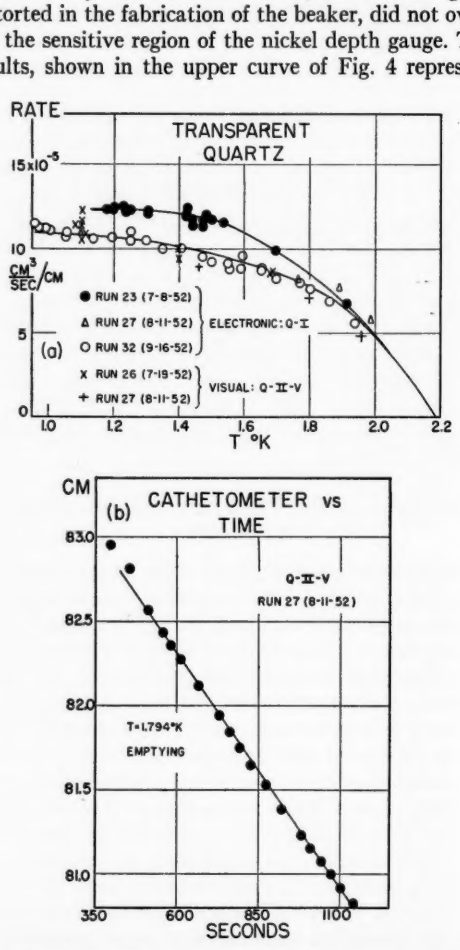


FIG. 5. (a) Transport rate over transparent quartz specimen Q-I, upper curve between 1.9 cm and 2.7 cm from the rim, lower curve between 1.5 cm and 3.6 cm from the rim. (b) Typical record of visual observation of rate of fall of helium meniscus in quartz capillary used in calculating rates shown in (a) for specimen Q-II-V.

rates averaged over distances from 1.5 cm to 3.6 cm from the beaker rim. Here too, the beaker was left undisturbed in the cryostat except for repumping between the two runs.

Quartz

Fused quartz was selected as a source of data for vitreous materials other than glass. Beakers were made from specimens of commercial tubing in each of the two available grades: transparent and translucent.

Transparent Quartz Specimen Q-I

Some of the results obtained with specimen Q-I, using a "three-section" nickel depth gauge, are given in the upper curve of Fig. 5(a), which represents the average rate between 1.9 cm and 2.7 cm from the beaker rim. Similar curves based on data associated with the other two calibrated regions (not shown) yielded rates 8 percent higher between 1.5 cm and 1.8 cm, and 8 percent lower in the region between 2.8 cm and 3.6 cm from the rim (see paper II).

Data obtained one and two months later are shown in the lower curve of Fig. 5(a). In these measurements the depth gauge which was used with specimen G-II was employed to provide data for comparison of transport over the two materials averaged over the same 1.5 cm to 3.6 cm distance from the rim.

Transparent Quartz Specimen Q-II-V

Visual observations of transport, averaged over similar heights of a carefully shielded clear quartz capillary (4 cm long, 0.1 cm bore), are also indicated in the lower curve of Fig. 5(a). A sample curve, indicating the course of cathetometer readings with time is shown in Fig. 5(b).

Translucent Quartz Specimen Q-III

The results of measurements similar to those on Q-I and G-II, covering film heights between 1.5 cm and 3.6 cm from the rim, are shown in Fig. 6. The beaker remained undisturbed in the cryostat, except for outgassing at room temperature between the two runs.

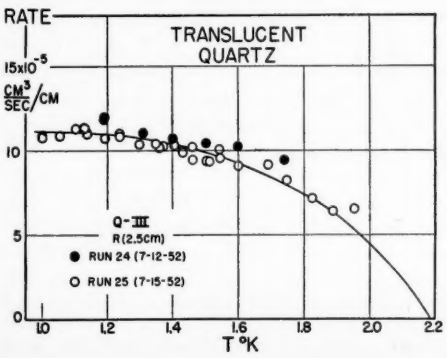


FIG. 6. Transport rates over translucent quartz specimen Q-III between 1.5 cm and 3.6 cm from the rim.

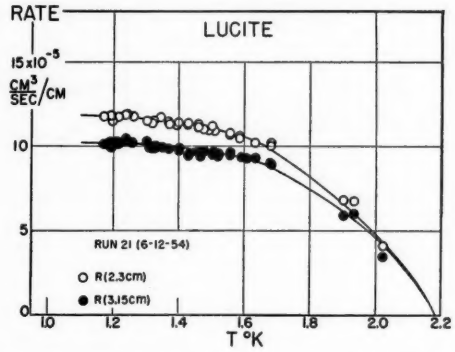


FIG. 7. Transport rates over Lucite, averaged between 1.9 cm and 2.7 cm from the rim (upper curve) and between 2.8 cm and 3.6 cm from the rim (lower curve).

Lucite

Chandrasekhar's report[18] on transport over methyl methacrylate polymers (Lucite and Perspex) provided new impetus for reinvestigation of a previous tentative result obtained in this laboratory.[13] Measurements were therefore undertaken using the original Lucite beaker with the addition of a new capacitor depth gauge. Rates averaged over distances from 1.9 cm to 2.7 cm and from 2.8 cm to 3.6 cm from the rim are shown respectively in the upper and lower curves of Fig. 7. Thus the current data do not duplicate the appearance of a maximum in the vicinity of 1.5°K. Instead, the results show a smooth variation with temperature, with departures in absolute value from the previous report[13] (omitting the peak) due primarily to the correction of the present data for the appreciable contraction of Lucite on cooling to helium temperatures, using expansion coefficients[29,30] which were unavailable at the time of the earlier work. Since interest in the present discussion is limited to a comparison of the magnitudes of the transport rates observed for various materials, a discussion of the peaks sometimes associated with rate-temperature characteristic curves is reserved for the

TABLE III. R_4: Average transport rate between 4.45 cm and 5.33 cm from beaker rim (midpoint at 4.89 cm).

Specimen	Run no.	Rate in cm³/cm-sec ×10⁸ 1.1°	1.3°	1.5°	1.7°	1.9°	2.1°
Al-I (*B*)	38, 48	9.7	9.6	9.1	7.9	5.9	2.5
Al-I (*B*)	39	11.3	11.2	10.8	9.4	7.2	3.5
Al-II (*B*)	55	9.2	9.0	8.7	7.8	6.1	2.6
Cu	59	10.1	10.0	9.7	8.6	5.8	2.0
Cu	66	9.4	9.3	9.0	8.0	5.7	2.0
Ni-Ag-I (*B*)	33–35, 37	12.1	12.0	11.9	10.4	7.2	2.6
Ni-Ag-II (*B*)	45, 46, 49	10.1	10.0	9.6	8.6	6.5	2.8
Ni-II (*R*)	53	12.9	12.5	11.7	10.3	7.4	2.9
Ag	58	8.6	8.4	8.0	7.3	5.9	2.5
S. St. (*B*)	41–44, 47	9.2	9.1	8.7	7.7	5.7	2.4
G-III	56	11.2	10.8	9.9	8.4	6.2	2.4

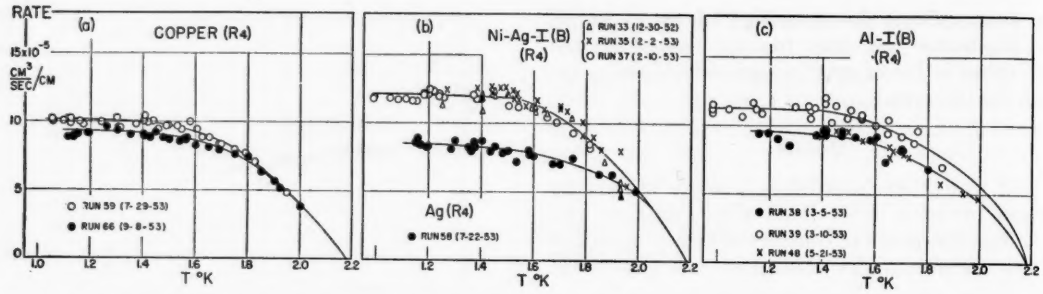


FIG. 8. Transport rates over various machined metal specimens: (a) copper, (b) nickel silver (upper curve), and silver (lower curve), (c) aluminum. In each case the transport rates were calculated for distances between 4.45 cm and 5.33 cm from the rim, corresponding to beaker region "4" in Fig. 2 and therefore designated as R_4.

last paper in this series (IV) which is devoted to a general study of the functional dependence on temperature of the transport rates.

The remainder of the data to be presented are summarized in Table III (where tabulation is facilitated by the adoption of the same geometry for all beakers listed). The designation R_4 corresponds to a measurement of transport in the beaker region occupied by the sensitive element labelled 4 in Fig. 2. In Table III, the numerical values were abstracted from smooth curves drawn through the experimental points. The original data for four of the nine specimens are shown in Fig. 8. Similar curves for the other five specimens will be found in paper II in connection with control experiments on surface finish (see part (a) of Figs. 2–6 of paper II). This allocation of curves is dictated by the desire to conserve space without losing sight of the fact that a complete display of the data is vital where so many conflicting reports exist. Reproducibility and scatter may also be judged from these curves.

DISCUSSION

Rates observed for all materials varied so little in comparison with the factor of three or more which one is led to expect from other investigations,[8,14,17] that the background variation prevents the resolution of any small systematic dependence on the chemical composition of the substrate which might exist. It is interesting to note that this does not preclude the existence of such a dependence in view of Schiff's calculation[39] of the size of the helium film-substrate interaction to be expected for a number of materials employed in this investigation. From consideration of the van der Waals and gravitational energies involved, Schiff found that the thickness, d, of the film at a height h above the free surface of the liquid would be given by

$$d = 4.3 \times 10^{-6}/h^{\frac{1}{3}} \text{ cm} \qquad (2)$$

for copper, whereas "the numerical value 4.3 for copper is replaced by 4.7 for silver" and "about 4 for glass."[39] Since the denominator of Eq. (2) remains the same for all substrates, the numerator is a direct measure of the expected film-substrate interaction. Thus the present observations, in combination with the calculations of Schiff argue against the assignment of a fundamental role to the results of a particular measurement of transport over glass.

The great similarity of the data for all these materials focuses attention on the desirability of ascertaining the contribution of surface roughness to those differences which were observed. Furthermore, there is sufficient evidence in Figs. 4–8 to indicate that a reliable comparison of the data averaged over different beaker heights also requires a careful consideration of the extent to which film height may influence the various data. Further discussion of these results will therefore be found in subsequent papers of this series.

ACKNOWLEDGMENTS

The authors wish to express their appreciation to Dr. J. G. Dash for assistance in the construction of the helium cryostat; to Albert Brown, Leonard Lesensky, and Robert D. Worley for valuable assistance during the course of the measurements; and to Professor Mark W. Zemansky for his active and stimulating interest in these experiments.

[39] L. I. Schiff, Phys. Rev. **59**, 839 (1941).

Electron Spin Resonance Absorption in Metals. I. Experimental

GEORGE FEHER* AND A. F. KIP
Department of Physics, University of California, Berkeley, California
(Received December 2, 1954)

The electron spin resonance absorption in metals was investigated at 300 and 9000 Mc/sec, in the temperature range from 4°K to 296°K. The metals investigated were: Li, Na, K, Be, Mg, Al, Pd, W. Resonances were observed in Li, Na, Be, for which accurate g values and relaxation times were obtained and compared with existing theories. K gave a signal only at 4°K and 300 Mc/sec, with too poor a signal-to-noise ratio to make accurate measurements. No resonances were observed in Mg, Al, Pd, W. The line shapes predicted by Dyson's theory for the resonance absorption in metals were checked experimentally for different diffusion times and relaxation times. A satisfactory agreement between the theory and experiment is obtained for both the normal and the anomalous skin effect region.

I. INTRODUCTION

THE first paramagnetic resonance absorption in metals due to conduction electrons was observed by Griswold, Kip, and Kittel.[1] At that time there was no theory to take the diffusion of the electrons properly into account and therefore exact g values and relaxation times could not be obtained. The purpose of this paper is to report on the extension of the work on electron spin resonance absorption in metals, making use of Dyson's theory (see following paper) on line shapes. With its aid accurate g shifts and relaxation times were obtained and are compared with existing theories.

II. THE EXPERIMENTAL SET-UP

Most of the experiments were performed at 300 megacycles per second. A block diagram of the experimental arrangement is shown in Fig. 1. The heart of the setup is a coaxial resonant cavity, which is one-half wavelength long and has a Q of about 1000–2000. The same cavity may be used at a frequency of 600 and 900 megacycles per second, in which case it will be a full wavelength or $\frac{3}{2}$ wavelength long. In either case it is used as a transmission cavity. The power is fed in and taken out by two identical coupling loops. Their length and orientation may be changed from outside to give optimum coupling conditions, which may be measured by means of the voltage standing wave ratio. Line stretchers are used for matching purposes.

The output of the cavity is rectified by means of a grid leak detector, which proved to be much less noisy than a crystal rectifier. The detector output feeds into a 6-kilocycle-per-second narrow-band amplifier. It has a Q of 85, maximum amplification of 120 db and an equivalent noise input voltage of 0.05 microvolt. It is followed by a lock-in detector which further cuts down the band width and hence the total amount of noise power coming through. Its output contains the desired signal and can be displayed either on a scope or Brown Recorder.

The magnetic field is obtained from a water-cooled solenoid (50 cm long) with an inner diameter of 6 cm. To improve on the uniformity of the field, compensating coils are wound on each end of the solenoid. This reduces the field variations to about one part in 10^5 for the central 10 cm of the solenoid. The magnetic field for a power input of 1.5 kw is 1000 oersteds, which is the maximum safe field to be used.

The dc magnetic field is modulated at a rate of six kc/sec by means of an auxiliary winding. Since the modulation is always small in comparison to the line width, the observed signal is proportional to the derivative of the absorption. The modulating coil is not wound along the entire length of the dc windings. This reduces the coupling between the two coils and hence reduces the shorting effect of the dc winding. To prevent the resonant cavity from acting as a shorted turn for the six kc/sec, it is provided with a longitudinal slot. Such a slot runs parallel to the rf currents in the cavity and therefore does not lower noticeably the Q of the cavity.

The dc field may be slowly varied by means of a motor driven rheostat. Unfortunately the variable contact resistance of the rheostat causes current fluctuations (and hence corresponding field fluctuations) of approximately one part in 10^4. Thus for narrow lines (of the order of 0.1 oersted) an alternate method of sweep was developed. It consists of a copper coil in an oil bath whose temperature may be either increased or decreased by means of a heating coil or water cooling system, respectively. As the temperature changes, the resistance of the coil changes and the current through the magnet varies. By this method a smooth variation of 10 percent of the current is easily attained. At 300 Mc/sec this corresponds to about 10 oersteds.

The magnetic field can be determined to an accuracy of one part in 10^5 by means of a standard resistor and potentiometer. The frequency is determined by means of a frequency meter to three parts in 10^5. This limits the accuracy with which the g values of the electrons in the investigated material may be determined. For broad lines an additional error is introduced in locating the exact resonance condition on the line. For those lines a 3-cm setup, having a higher dc field, is better suited for

*Now at Bell Telephone Laboratories, Murray Hill, New Jersey.

[1] Griswold, Kip, and Kittel, Phys. Rev. 88, 951 (1952).

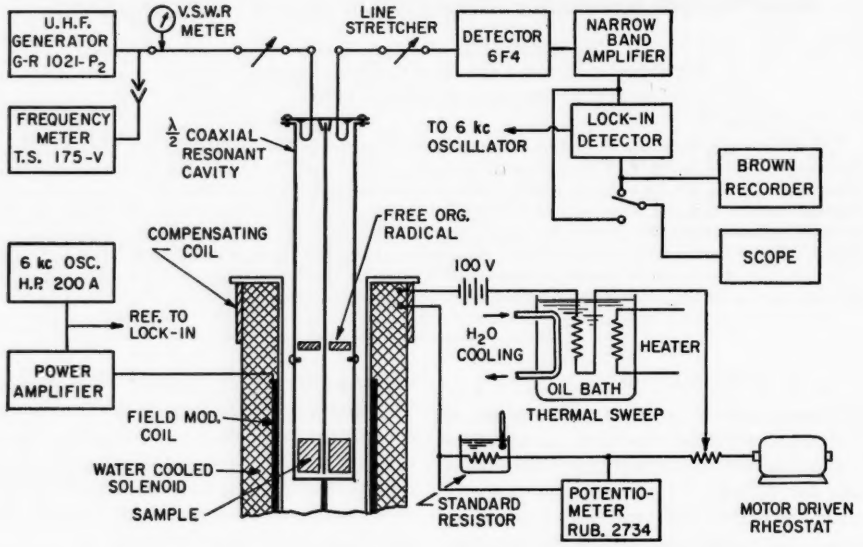


FIG. 1. Experimental setup for electron spin resonance absorption at 300 Mc/sec.

exact g determinations and was used by us for this purpose.

For the absolute measurement of susceptibilities, and for general calibration purposes a known amount of an organic free radical (diphenyl picryl hydrazyl) is placed in the cavity simultaneously with the sample. The position of the free radical sample is such that it sees a lower rf field, lower dc magnetic field and a lower modulation field. The reason for choosing its position in this way is as follows: it prevents the free radical absorption from falling on top of the investigated lines, which almost invariably are located around $g=2.00$. The different dc magnetic field displaces it sufficiently to avoid any interference. (See Fig. 9.) Most lines investigated have broader lines than the diphenyl picryl hydrazyl (which, depending on the temperature, has a width of approximately one to two oersteds between inflection points). Thus for a proper sample modulation, the free radical would be overmodulated. The fact that it is located in a weaker rf magnetic field makes it a valuable reference for saturation measurements where rf fields of the order of oersteds may have to be used.

In order to cool the sample to below room temperature, two coaxial Dewars are used. Into the inner one the cavity is suspended by two coaxial lines (input and output) whose inner and outer conductors are made out of brass to reduce the heat loss by conduction.

The range between liquid nitrogen (77°K) and room temperature is covered by filling the space between the two Dewars with liquid nitrogen and mounting a heater in thermal contact with the cavity. By adjusting the power input to the heater, any equilibrium temperature within the above range may be obtained and kept constant to a fraction of a degree. In order to prevent the condensation of oxygen below 90°K, the inner Dewar in which the cavity is situated is made vacuum tight and filled with helium gas.

The temperature range between 4°K, and 77°K, is covered by mounting a brass rod at the end of the cavity. The temperature of the cavity will then be determined by the position of the liquid-helium level on the brass rod. In order to avoid temperature changes while sweeping through a line, parts of the brass rods are thermally "shorted out" by copper blocks. Since the thermal conductivity of copper at 4°K is about one hundred times higher than that of brass, no appreciable difference in temperature occurs while the helium is in contact with the copper blocks. This method works only if the rate of evaporation is low. If this is not the case, the helium vapor rather than the brass rods will be effective in cooling the cavity. The rate of loss of liquid helium without any rf in our system is 30 cc/hour which corresponds to about 25 mW. Since the rf powers ordinarily used are of the same order, an improvement of the Dewar design would not result in an appreciably lower rate of helium loss. Incidentally, a measurement of the loaded and unloaded Q of the cavity, together with the rate of helium evaporation, may be used to calculate the rf magnetic field in the cavity. This is an important quantity for saturation measurements.

It was found that when the cavity is filled with liquid helium, the noise output increases greatly. This extra noise source was attributed to bouncing helium bubbles within the cavity. In order to eliminate this trouble, the entire cavity is filled with Styrofoam, a low-loss material with a dielectric constant of 1.0.

The temperatures are read with an AgAu-AuCo

thermocouple, which, unlike most other thermocouples, is still useful in the 4°K range. At this temperature it has a slope of approximately five microvolts per degree. The thermocouple is soldered to the wall of the cavity at the place where the sample is situated.

The sensitivity of the 300-Mc equipment at room temperature, set by the random noise fluctuations, corresponds to about 10^{16} free electron spins having a line width of one oersted. A curious sensitivity limit for broad lines is set by what is loosely called cavity effects. Those are reproducible lines of several oersteds width and occur at different g-values. They probably are the result of impurity inclusions in the walls of the cavity. Since the occurrence of the peaks depends on the previous history of the magnetic field, they are presumably ferromagnetic inclusions. A more careful silver plating job probably would eliminate this trouble.

Let us consider now how the sensitivity changes on going from 300 Mc/sec (at which frequency most of this work was done) to 9000 Mc/sec, which is the more conventional frequency for electron spin resonance experiments. The higher fields which are used with higher frequencies give rise to a larger difference in population of spin states, and hence to a larger signal. However, for metal samples large compared to a skin depth, the effective volume of the sample goes down with increasing frequency (because of the skin depth), so that for a constant line width with comparable Q's and filling factors the fractional power absorbed at resonance

$$\Delta P_C/P_C \cong (\omega)^{\frac{1}{2}}.$$

This would indicate that the expected sensitivity at 9000 Mc/sec should be about 6 times better than at 300 Mc/sec. However, the crystal detectors used at microwave frequencies are much noisier than the vacuum tubes used at uhf frequencies, so that we get about equal sensitivities as far as metals are concerned.

The advantage of the lower frequency set-up was the homogeneity and stability of the magnetic field, which made it possible to investigate very narrow lines (e.g., sodium at 4°K). On the other hand, lines broader than 100 oersteds (the field at 300 Mc/sec) cannot be investigated profitably at 300 Mc/sec.

III. PREPARATION OF SAMPLES

The samples investigated were of two forms. Either their dimensions were small in comparison to a skin depth at 300 Mc/sec (i.e., thin films or dispersions of particles), or large in comparison to a skin depth (i.e., sheets or large particles).

For all metals except the alkalis, the sample had dimensions large in comparison to a skin depth. When particles were used, they were imbedded in paraffin and molded into the desired shape. When metal plates were used, they were sandwiched between polystyrene sheets in order to expose more sample volume to the rf field. In this case it is important to insert the metal plates into the cavity in such a way that the E field is perpendicular to the metal surface. This reduces the losses due to the sample.

Small particles of the alkalis were prepared by ultrasonerating the bulk metal in mineral oil at a frequency of 700 kc/sec. The average particle size of such a dispersion was $\sim 15\mu$. In order to obtain a dispersion with all particles smaller than a skin depth, the sample was centrifuged at $2000g$ for several hours which sedimented out the large particles. (For the experimental result on such a centrifuged dispersion, see Fig. 8.)

Sheets of the alkali metals were prepared by rolling out the bulk material in molten paraffin. Mineral oil was found to react with the metal to form free radicals which obscured the metallic resonance at low temperatures (see Sec. IV, A-2) and therefore was substituted by the molten paraffin. An alternate way of making thick sheets of Li and Na is to put the bulk metal between two sheets of Parafilm (a thin strengthened sheet of paraffin, commercially available from the Central Scientific Company) and roll it out on a regular sheet metal roll. This is a much faster method, but the samples have to be used within a few days since they are more likely to react in air.

Some alkali films (up to 0.15-mm thickness) were also evaporated in vacuum on thin mica sheets. Before exposing them to air, a thin protective layer of paraffin was evaporated onto them. The main purpose of this technique was to purify the metals. (See Sec. IV, C.) In the case of thin films this also provided the sample for the determination of the absolute value of the susceptibility.

During the writing of this paper some work performed with R. A. Levy indicates that when alkali metals are dissolved in ammonia, they precipitate out as metallic particles below the freezing point of ammonia (~190°K). This seems to be a promising way of preparing samples, especially of the more reactive alkalis.

IV. EXPERIMENTAL RESULTS AND THEIR COMPARISON WITH EXISTING THEORIES

A. Line Shapes in the Region of the Normal Skin Effect

1. Dyson's Theory

The references to Dyson's formulas in the following paper will be prefixed by *D*. Dyson treats the problem of electron diffusion in connection with the paramagnetic resonance in metals and obtains closed expressions for different experimental conditions. We will recapitulate briefly (in a slightly simplified form) some of his main results which are going to be used throughout this paper.

The main parameters in Dyson's theory are: T_D = time it takes an electron to diffuse through the skin depth; T_1 = electron spin relaxation time; (for metals $T_1 = T_2$); T_T = time it takes the electron to traverse the sample. The simplest case is:

$$T_T \ll T_D;$$

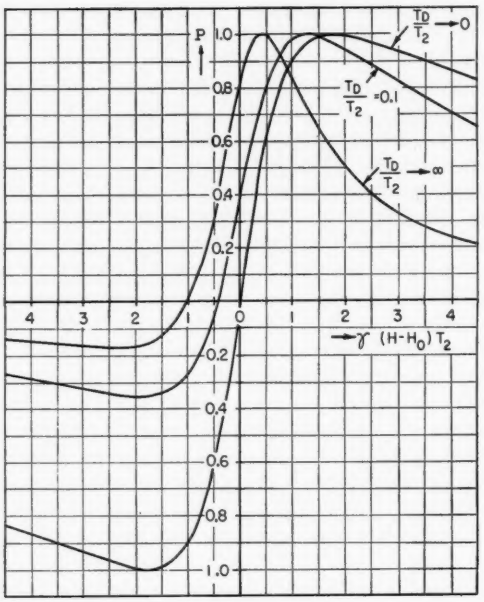


FIG. 2. Power absorption due to electron spin resonance in thick metal plates for different ratios of diffusion time T_D to relaxation time T_2.

i.e., a film (or particle) of thickness small in comparison to the skin depth. Under those conditions the power absorbed (D-78) is

$$P=\frac{\omega H_1^2}{4}(V)\omega_0\chi_0 T_2\frac{1}{1+(\omega-\omega_0)^2T_2^2}, \tag{3.1}$$

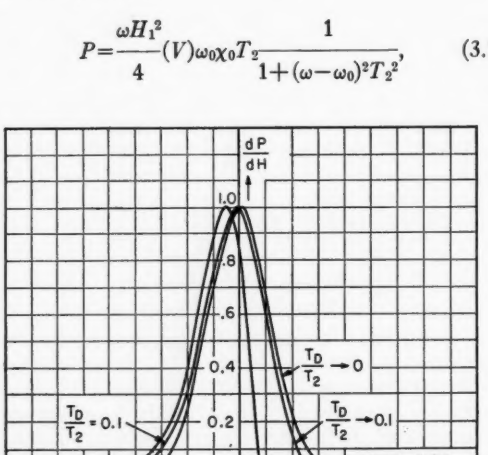


FIG. 3. Derivative of the power absorption due to electron spin resonance in thick metal plates for different ratios of diffusion time T_D to relaxation time T_2.

and the experimentally observed quantity is

$$\frac{dP}{d\omega}=-\frac{\omega H_1^2}{4}(V)\omega_0\chi_0 T_2\frac{2(\omega-\omega_0)T_2^2}{[1+(\omega-\omega_0)^2T_2^2]^2}, \tag{3.2}$$

where V = volume of the sample, H_1 = amplitude of the linearly polarized magnetic rf field, χ = paramagnetic part of the static susceptibility, and ω_0 = the resonant frequency satisfying the condition $\hbar\omega_0=g\beta H$.

As expected, this result is entirely independent of diffusion and gives a symmetrical Lorentz line with the natural half-width $1/T_2$.

A more important case is:

$$T_T\gg T_D, \quad T_T\gg T_2.$$

This condition is always fulfilled for sufficiently thick samples. The general result for an arbitrary ratio of T_D/T_2 is (D-82):

$$P=-\left[\frac{\omega H_1^2}{4}(\delta A)\omega_0\chi_0 T_2\right]\frac{T_D}{2T_2}\left\{\frac{R^4(x^2-1)+1-2R^2x}{[(R^2x-1)^2+R^4]^2}\times\left[\frac{2\xi}{R(1+x^2)^{\frac{1}{2}}}+R^2(x+1)-3\right]+\frac{2R^2-2xR^4}{[(R^2x-1)^2+R^4]^2}\times\left[\frac{2\eta}{R(1+x^2)^{\frac{1}{2}}}+R^2(x-1)-3\right]\right\}, \tag{3.3}$$

where $x=(\omega-\omega_0)T_2$, $\xi=(\text{sign of } x)[(1+x^2)^{\frac{1}{2}}-1]^{\frac{1}{2}}$, $\eta=[(1+x^2)^{\frac{1}{2}}+1]^{\frac{1}{2}}$, $R=(T_D/T_2)^{\frac{1}{2}}$, $\delta=c/(2\pi\sigma\omega)^{\frac{1}{2}}$ = classical skin depth, and A = area of surface.

Equation (3.3) is plotted in Fig. 2 for different ratios of T_D/T_2. The experimentally observed derivative of P with respect to H was taken graphically from plots of Eq. (3.3) or computed for the cases $T_D/T_2\ll 1$; $T_D/T_2\to\infty$ for which asymptotic expressions were derived. They are plotted in Fig. 3.

Since those limiting cases are much easier to handle than Eq. (3.3) it seemed worthwhile to list them below.

(a) $T_T\gg T_D$; $T_T\gg T_2$; $T_D/T_2\to 0$.

This condition applies to metals of high conductivity (low temperature) and narrow lines (D-88).

$$P\cong-\left[\frac{\omega H_1^2}{4}(\delta A)\omega_0\chi_0 T_2\right]\left\{\left(\frac{T_D}{T_2}\right)^{\frac{1}{2}}\times\frac{[(1+x^2)^{\frac{1}{2}}-1]^{\frac{1}{2}}}{(1+x^2)^{\frac{1}{2}}}(\text{sign of } x)\right\}, \tag{3.4}$$

$$\frac{dP}{d\omega}=-\left[\frac{\omega H_1^2}{4}(\delta A)\omega_0\chi_0 T_2\right]\left\{\frac{T_2}{2}\left(\frac{T_D}{T_2}\right)^{\frac{1}{2}}\times\frac{[2-(1+x^2)^{\frac{1}{2}}][(1+x^2)^{\frac{1}{2}}+1]^{\frac{1}{2}}}{(1+x^2)^{\frac{1}{2}}}\right\}. \tag{3.5}$$

(b) $T_T\gg T_D$; $T_T\gg T_2$; $T_D/T_2\ll 1$.

This case is similar to the preceding one except that we keep terms of the order T_D/T_2. Thus we have to add to Eq. (3.5) a correction term.

$$\frac{dP}{d\omega}=\text{Eq. (3.5)}-\left[\frac{\omega H_1^2}{4}(\delta A)\omega_0\chi_0 T_2 \times\left\{T_2\left(\frac{T_D}{T_2}\right)^{\frac{1}{2}}\frac{[(1+x^2)^{\frac{1}{2}}-1]^{\frac{1}{2}}}{(1+x^2)^{\frac{1}{2}}}\right\}\right]. \quad (3.6)$$

(c) $T_T \gg T_D$; $T_T \gg T_2$; $T_D/T_2 \gg 1$.

This is again the thick-plate case with slowly diffusing magnetic dipoles. This case is well known from nuclear resonance in metals,[2] since the nuclei are almost completely stationary. This condition would also prevail in the case of paramagnetic impurities distributed throughout the volume of the metal. For these conditions we have: (D-84)

$$P=\left[\frac{\omega H_1^2}{4}(\delta A)\omega_0\chi_0 T_2\right]\left\{\left(\frac{1}{2}\right)\frac{1-T_2(\omega-\omega_0)}{1+T_2^2(\omega-\omega_0)^2}\right\}, \quad (3.7)$$

$$\frac{dP}{d\omega}=\left[\frac{\omega H_1^2}{4}(\delta A)\omega_0\chi_0 T_2 \times\left\{\left(\frac{T_2}{2}\right)\frac{T_2^2(\omega-\omega_0)^2-2T_2(\omega-\omega_0)-1}{[1+T_2^2(\omega-\omega_0)^2]^2}\right\}\right]. \quad (3.8)$$

All of the expressions (3.1) to (3.8) are only approximations to Dyson's formulas which include another term equal to the expressions themselves, but having $(\omega-\omega_0)$ replaced by $(\omega+\omega_0)$. (Note: ω_0 and H are freely interchanged in all the discussions, because of their relation: $\hbar\omega_0=g\beta H$, i.e., $\omega_0=\gamma H$.) For reasonably narrow lines this term may be neglected near a resonance. Physically this term arises from the use of a linearly polarized magnetic rf field, which is a superposition of two circularly polarized fields. One of them, having the same sense of rotation as the precession of the electron, is much more effective than the other in flipping the spins.

Another effect which is not included in expressions (3.1) to (3.8) is that of surface relaxations, which are important whenever there exist strong spin-dependent forces during a collision of the electron with the surface. Dyson treats this modification and shows that in the case of thin films or small particles one expects a much more marked effect than in the thick cases. From the experimental results, to be discussed in the next section, we can conclude that for the case of thick alkali metals this effect may be neglected.

Having a theory for the line shape, one may now proceed to extract the desired information from it.

Figures 4 and 5 show the point on the curve at which the resonance condition is satisfied for different ratios of T_D/T_2. This enables one to determine accurate g values.

[2] N. Bloembergen, J. Appl. Phys. **23**, 1379 (1952).

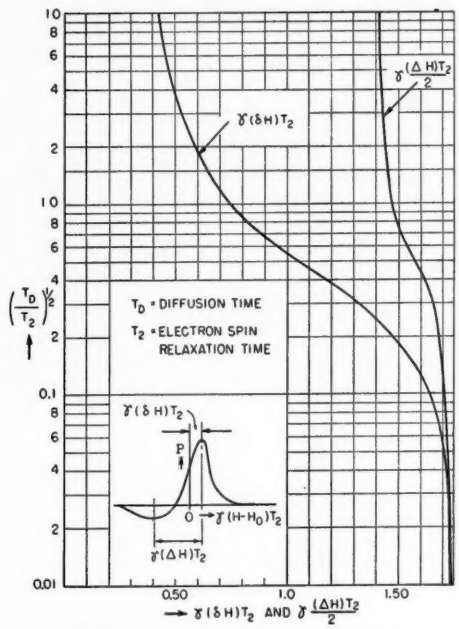


FIG. 4. $\gamma(\delta H)T_2$ and $\gamma(\Delta H)T_2/2$ *vs* $(T_D/T_2)^{\frac{1}{2}}$ for the power absorption due to electron spin resonance in thick metal plates.

They also show the way in which the half-power absorption points may be used to determine the relaxation time T_2.

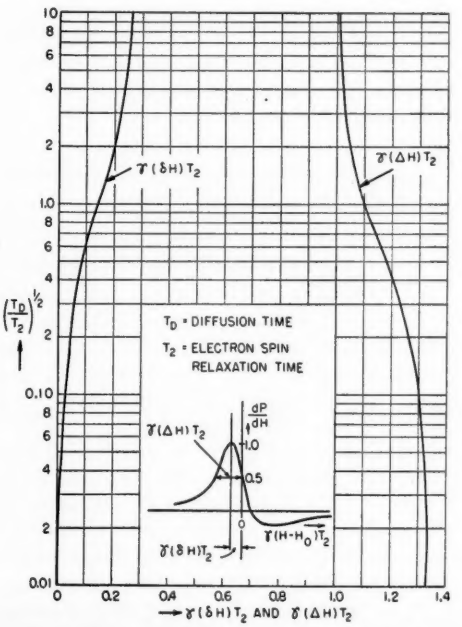


FIG. 5. $\gamma(\Delta H)T_2$ and $\gamma(\delta H)T_2$ *vs* $(T_D/T_2)^{\frac{1}{2}}$ for the derivative of the power absorption due to electron spin resonance in thick metal plates.

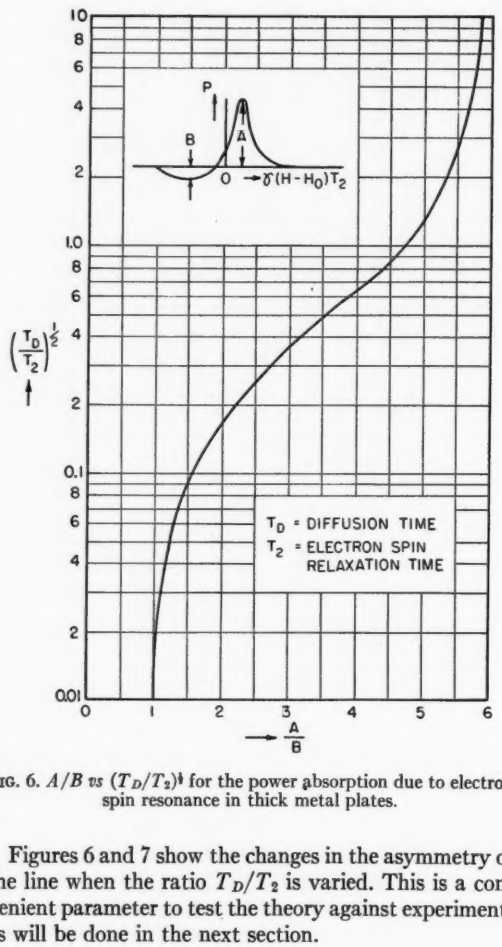


FIG. 6. A/B *vs* $(T_D/T_2)^{\frac{1}{2}}$ for the power absorption due to electron spin resonance in thick metal plates.

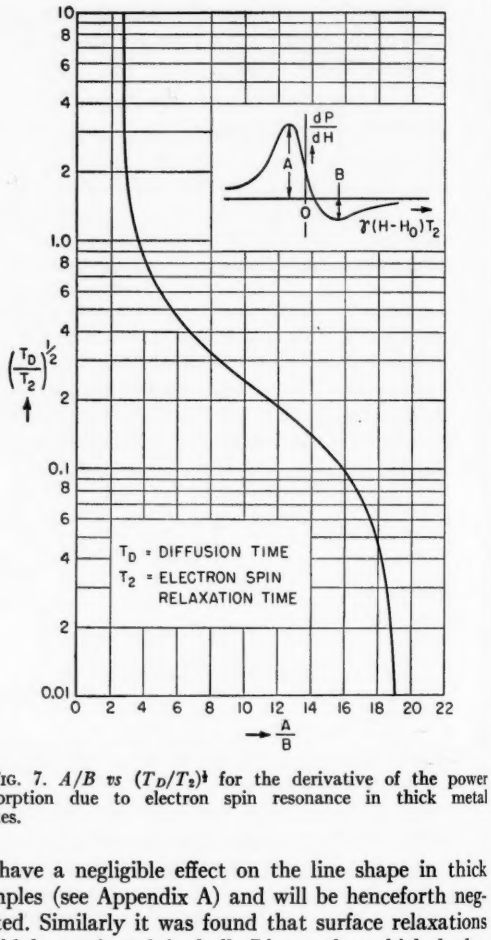


FIG. 7. A/B *vs* $(T_D/T_2)^{\frac{1}{2}}$ for the derivative of the power absorption due to electron spin resonance in thick metal plates.

Figures 6 and 7 show the changes in the asymmetry of the line when the ratio T_D/T_2 is varied. This is a convenient parameter to test the theory against experiment, as will be done in the next section.

2. *Comparison of Dyson's Theory with Experiment*

In order to determine to what extent surface relaxations are important, the line widths in dispersed sodium and lithium were compared with those in bulk samples (for the preparation techniques, see Sec. III). The particle size in all dispersions was smaller than 5μ, which is less than the skin depth at 320 Mc/sec. At $T=296°$K and $f=300$ Mc/sec the skin depth for Na is 6μ and for Li, 8μ. The theoretical expression is given for this case by Eq. (3.2). It is compared in Fig. 8 with the experimental curve obtained from a sodium dispersion with which it agrees very well. This confirms Dyson's assumption that the relaxation proceeds exponentially. From the curve we find a total relaxation time $T_2=6\times10^{-9}$ sec, as compared with the relaxation time in the bulk metal of $T_2=9\times10^{-9}$ sec. This discrepancy is within the variation of the relaxation times observed in different samples. But even if this difference were genuinely due to a surface phenomenon, it would lead to a surface relaxation time of 2×10^{-8}, which can be shown to have a negligible effect on the line shape in thick samples (see Appendix A) and will be henceforth neglected. Similarly it was found that surface relaxations could be neglected in bulk Li samples which had a relaxation time of 2×10^{-8} sec.

We proceed now to the more important case of thick plates. Figure 9 shows the experimental results in Li for different ratios of T_D/T_2. We see that as T_D/T_2 decreases, the sharp minimum of the derivative of the absorption disappears. Comparing the curves with Fig. 3, we see immediately the quantitative agreement with Dyson's theory.

The resistivity data which were used to calculate the diffusion time were taken from the work of MacDonald and Mendelssohn.[8] The Li used was obtained from the Lithium Corporation of America and was 99.9 percent pure; the sodium was obtained from A. D. MacKay Company and was 99.99 percent pure.

For a more quantitative test of Dyson's theory, the ratio between the maximum and minimum of the derivative of the absorption *vs* field was taken. This ratio A/B, according to Dyson's theory, should change from 2.7 to 19 over the entire range of T_D/T_2 (see Fig. 7). It

[8] D. K. C. MacDonald and K. Mendelssohn, Proc. Roy. Soc. (London) **A202**, 103 (1950).

is a convenient parameter since it may be varied easily by changing the temperature of the sample.

Figure 10 shows the comparison between theory and experiment. It is seen that most of the experimental points lie to the left of the curve, indicating that the diffusion time of the electrons is longer than that calculated from the resistivity data. This could be explained under the assumption that surface irregularities impede the diffusion of the electrons into the metal and thus lengthen T_D. This assumption seems to be born out by a sodium sample which had an exceptionally smooth surface and lies very near the theoretical curve. It was prepared by pouring molten sodium under vacuum conditions into the space between two smooth concentric glass cylinders. The Be sample which had a shiny metallic surface lies also on the theoretical curve.

We may thus conclude that the agreement between Dyson's theory and experimental results on thick plates is satisfactory.

In this connection it should be noted that the proper value of the experimentally observed parameter A/B may always be taken as conclusive evidence that the absorption is due to diffusing electrons. By a careful determination of this ratio, one can find the diffusion time of the electrons. The theoretical expression for the diffusion time under the assumption that the mean free path Λ is small in comparison to a skin depth and

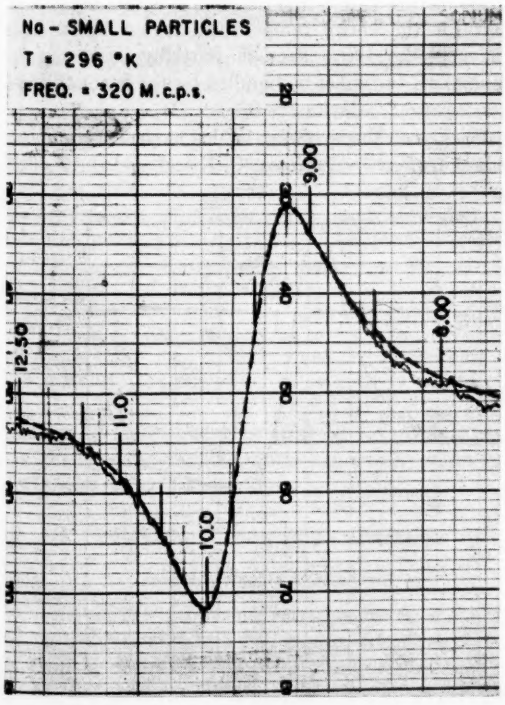
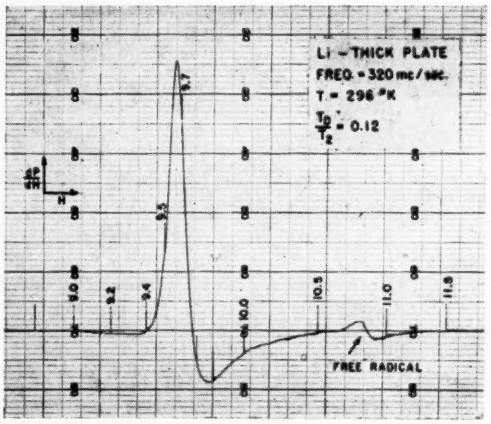
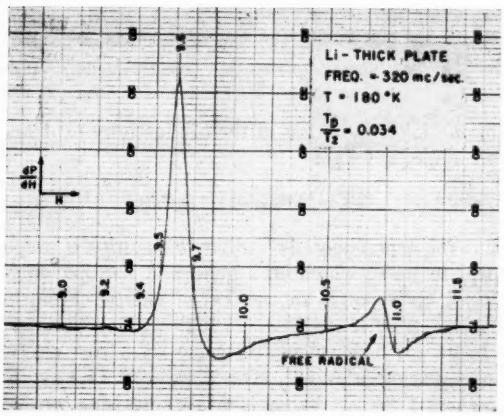


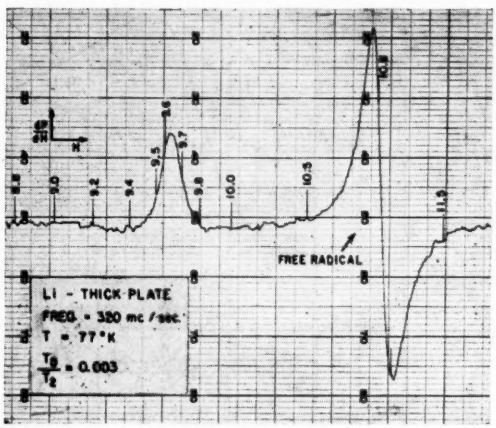
FIG. 8. Electron spin resonance absorption in small sodium particles. Diameter of particles smaller than 5 microns.



(a)



(b)



(c)

FIG. 9 (a, b, c). Electron spin resonance in thick plates of lithium for different ratios of T_D/T_2. Markers are proportional to the magnetic field. 1 unit = 11.77 oersteds. Compare the shape with the theoretical curves of Fig. 3.

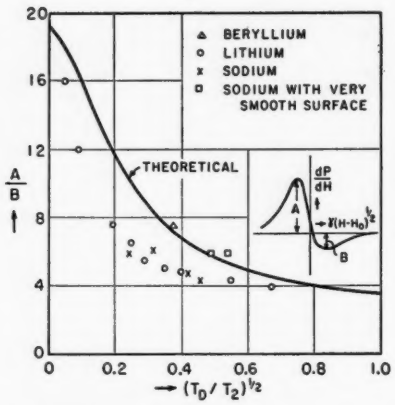


FIG. 10. Comparison of the asymmetry parameter A/B with theory for different ratios of T_D/T_2. The discrepancy is believed to be due to surface irregularities.

that all the electrons have the same velocity v is given by (D-11)

$$T_D=\tfrac{3}{2}(\delta^2/v\Lambda).$$

Eliminating Λ by means of the conductivity expression $\sigma=Ne^2\Lambda/m^*v$, we get

$$T_D=\tfrac{3}{2}(\delta^2Ne^2/\sigma)(1/m^*v^2).$$

Thus the experimentally measured quantity is proportional to $\frac{1}{2}m^*v^2$, which for a quadratic isotropic energy surface corresponds to the Fermi energy. From the observed ratio of A/B in beryllium we find that $\frac{1}{2}m^*v^2=16\pm2$ ev. The calculated value on a free electron model, taking $m^*/m=1$, gives a value of 14.3 ev. The closeness of the two values is rather surprising considering the crudeness of a free electron picture for beryllium.

Another importance of the parameter A/B lies in the fact that an undesired resonance due to impurities may be spotted immediately, since stationary paramagnetic impurities at the surface of the metal would give $A/B=1$; if they are distributed throughout the volume $A/B=2.7$ for a Lorentzian line and $A/B=2.0$ for a Gaussian line. The ratios 1 and 2.0 can never be due to conduction electrons in thick plates, whereas the ratio 2.7 could only occur in the limiting case of very slow diffusion and broad lines never observed in practice. This is a definite advantage of working with thick plates rather than small particles. In the latter case $A/B=1$, the same as for any kind of impurities. This can be very misleading (and did indeed mislead us for a while), especially when working with alkali metals. They form free radicals with mineral oil which at liquid helium temperature can completely obscure the metallic resonance (due to their $1/T$ dependence). Their line width is several oersteds at liquid helium temperature. Waxes which are solid at room temperature are much less troublesome in this respect. We found that the best of them comes under the trade name of "Parowax."

B. Line Shapes in the Region of the Anomalous Skin Effect

The expression for the classical skin depth is only valid when the mean free path Λ of the electrons in the metal is small in comparison to the skin depth δ. For the alkali metals at 300 mc, δ equals Λ at about 40°K, so that for the entire temperature region below liquid nitrogen the skin effect is partially or completely anomalous and must be handled by the exact theory of Reuter and Sondheimer.[4]

The Dyson theory was extended to the case of anomalous skin effect by Professor C. Kittel, and the resulting

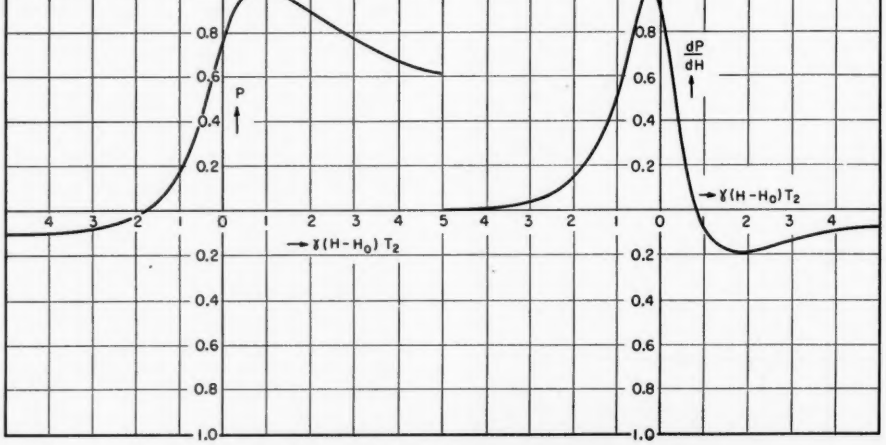


FIG. 11. Theoretical curves for the electron spin resonance absorption in the region of the completely anomalous skin effect.

[4] G. E. H. Reuter and E. H. Sondheimer, Proc. Roy. Soc (London) **A195**, 336 (1948).

line shape is given by Eqs. (D-26) and (D-40). The extended theory is valid only under the assumption that the diffusion time D is short in comparison to the relaxation time. This assumption is valid for all the metals in which resonances were observed at low temperatures.

The result for the power absorbed, P, when sweeping through a line is:

$$P=C\left\{\frac{(B^2-A^2)[(1+x^2)^{\frac{1}{2}}+1]^{\frac{1}{2}}}{(1+x^2)^{\frac{1}{2}}}+\frac{2AB[(1+x^2)^{\frac{1}{2}}-1]^{\frac{1}{2}}}{(1+x^2)^{\frac{1}{2}}}(\text{sign of } x)\right\}. \quad (3.9)$$

The experimentally measured quantity is the derivative of the power absorbed with respect to the field:

$$\frac{dP}{dH}=C'\left\{\frac{B^2-A^2}{2}\times\frac{[2+(1+x^2)^{\frac{1}{2}}][(1+x^2)^{\frac{1}{2}}-1]^{\frac{1}{2}}}{(1+x^2)^{\frac{1}{2}}}(\text{sign of } x)+AB\frac{[(1+x^2)^{\frac{1}{2}}-2][(1+x^2)^{\frac{1}{2}}+1]}{(1+x^2)^{\frac{1}{2}}}\right\}, \quad (3.10)$$

where

$$C=\tfrac{1}{4}\omega H_1^2(\Lambda A)\omega_0\chi_0(6T_2\tau)^{\frac{1}{2}}(C^2/4\pi\omega\Lambda)^2,$$

$$C'=\gamma T_2 C.$$

A and B are defined by the complex impedance $Z=A+iB$. They are functions of the classical skin depth and the mean free path. Their ratio may be conveniently read off from Reuter and Sondheimer[4] (Fig. 1). For the region of the completely anomalous skin effect, $B/A=\sqrt{3}$. For this case, expressions (3.9) and (3.10) are plotted in Fig. 11. Taking the resistivity data of MacDonald and Mendelssohn[3] for pure sodium at 4°K, one obtains the parameter, α, defined by Reuter and Sondheimer,

$$\alpha=\tfrac{3}{2}(\Lambda/\delta_0)^2\cong 4\times 10^5 \text{ at } 300 \text{ Mc/sec.}$$

This is well in the region of the completely anomalous skin effect.

Figure 12 shows the agreement between the experimental curve and the theoretical expression for sodium at 4°K and 300 Mc/sec.

C. Relaxation Times *vs* Temperature

1. Experimental Results

The metals in which resonances were observed may be conveniently divided into two groups, according to the temperature dependence of their relaxation times. In Li and Be, the relaxation time is temperature independent whereas in Na and K it varies with temperature. The signal to noise ratio for the K sample was too poor to get accurate data. Since the line shape in K could not be checked against Dyson's theoretical curves, there exists a slight doubt as to whether the signal was due to conduction electrons. The measurements were made at 300 Mc/sec.

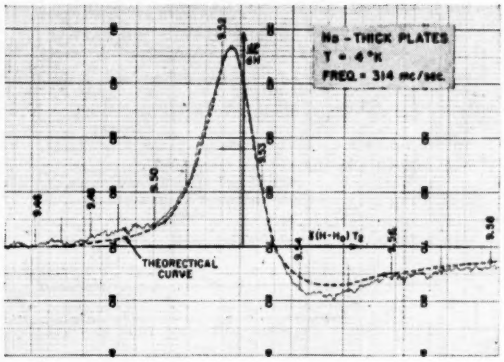


FIG. 12. Comparison between theory and experiment of the electron spin resonance absorption in the region of the completely anomalous skin effect.

Na, K.—The relaxation time *vs* temperature for sodium is plotted in Fig. 13, from which we see that $T_2\sim 1/T$.

For K we observed only a signal at 4°K which gave an estimated

$$T_2=5\times 10^{-9} \text{ sec.}$$

If the line width had stayed constant on heating up the sample, the amplitude of the observed signal should have increased (because of the increase in skin depth). Since, however, no resonance was observed above 4°K,

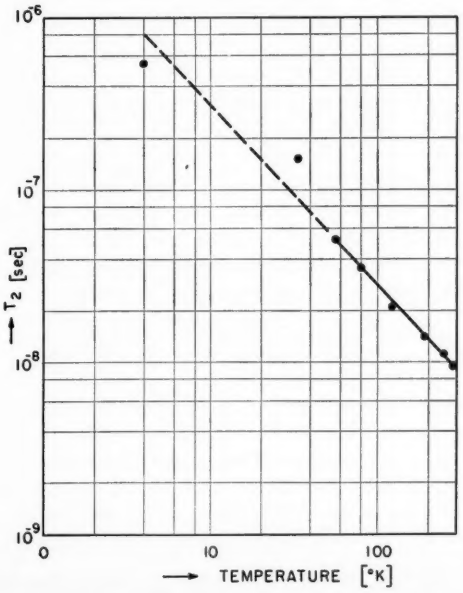


FIG. 13. Electron spin relaxation time T_2 *vs* temperature in sodium.

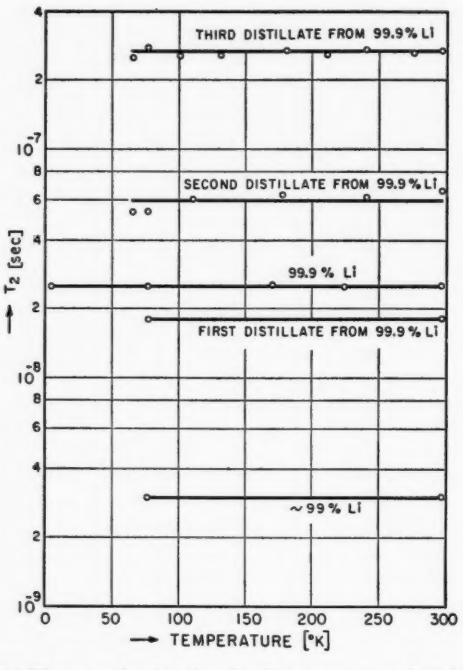


FIG. 14. Electron spin relaxation time T_2 *vs* temperature for lithium samples of different purity.

it must be concluded that the line width increases (T_2 decreases) with temperature.

Li, Be.—The relaxation time in both metals stayed constant in the temperature range between 4°K and 296°K.

From Fig. 14, we see that the relaxation time in lithium increases from 3×10^{-9} to 2.5×10^{-8} sec, upon purification of the sample from 99.0 percent to 99.9 percent. This strongly suggests that the relaxation is associated with impurities, rather than the metal proper. To prove this point even more conclusively, several distillates were taken from the 99.9 percent pure material (for the techniques see Sec. III). The first distillate had a shorter relaxation time ($T_2=1.8\times10^{-8}$ sec) than the subsequent ones ($T_2=6\times10^{-8}$ sec and 2.7×10^{-7} sec respectively). This result, besides proving the point made earlier, shows that the impurities responsible for the relaxation mechanism have a higher vapor pressure than lithium. It would be very desirable to purify the lithium further in order to get the "true" relaxation time.

Several different samples of beryllium were tried, but only one, supplied by the French Company, Pechiney (their U. S. representative is International Selling Corporation, 122 East 42nd Street, New York 17, N. Y.), and quoted to be 99.5 percent pure, gave a resonance signal. The relaxation time was temperature independent and had a value of

$$T_2=2.0\times10^{-8}\text{ sec.}$$

The temperature independence of T_2 suggests again an impurity scattering mechanism. The samples which showed no resonance had presumably a higher impurity concentration. (These were supplied by Brush Beryllium Company, and Fairmont Chemical Company. We are also indebted to Professor E. Segrè for supplying us with beryllium sheets.) This would broaden the line and consequently reduce the signal below the noise level. This could also explain the failure of Gutowsky and Frank[5] to observe a resonance in beryllium.

2. *Comparison of Experimental Results with Overhauser's and Elliott's Theory*

The paramagnetic relaxation in metals was treated by Overhauser.[6] His shortest relaxation time is 8×10^{-7} which is two orders longer than the experimentally observed value in sodium. This particular mechanism has also a field dependence which should increase the relaxation time by 40 percent when going from 300 Mc/sec (~100 oersteds) to 9000 Mc/sec (~3000 oersteds). The experimentally determined ratio of T_2 for these two frequencies for the same sodium sample was found to be (T_2 at 100 oersteds)/(T_2 at 3000 oersteds) $=1.05\pm0.05$. This result indicates that the above relaxation mechanism is not merely off by a proportionality constant.

It is of interest to note that the interaction of the spins with longitudinal phonons gives the observed temperature dependence in Na but is off by 4 orders. The large discrepancy between Overhauser theory and the experimental results is probably due to his assumption of plane waves for the electronic wave functions.

Before leaving his theory, one result is worth noticing. The spin-spin interaction, which would broaden the line without providing a spin lattice relaxation mechanism, is completely negligible. This is the theoretical justification for setting $T_1=T_2$. The experimental proof was given by Slichter[6] who demonstrated the validity of this relation from saturation measurements.

Another possible explanation for the spin relaxation mechanism has been suggested by Elliott,[7] who considered the effect of spin-orbit coupling on the relaxation time. Since the interaction of the lattice with the motion of the electrons is very large in comparison with the direct interaction of the spins with the lattice, a very small coupling of the spins to their orbits will provide a very effective relaxation mechanism. His result for the spin relaxation time T_1 is [see Eq. (59), reference 7]

$$T_1=\alpha[\tau_R/(\Delta g)^2], \tag{3.13}$$

where α=numerical factor smaller than 1; if one assumes the simple Debye theory for the lattice waves, $\alpha\cong1/30$; τ_R=the usual relaxation time obtained from resistivity measurements; $\Delta g=(g-2.0023)$ deviation of

[5] H. S. Gutowsky and P. J. Frank, Phys. Rev. **94**, 1067 (1954).
[6] A. W. Overhauser, Phys. Rev. **89**, 689 (1953).
[7] R. J. Elliott, Phys. Rev. **96**, 266, 280 (1954).

the electronic g value from the free electron value; this is a measure of the amount of spin-orbit coupling.

For sodium, $\tau_R=3\times10^{-14}$, $\Delta g=8\times10^{-4}$ (see next section), we get for T_1 a value of 2×10^{-9}. Comparing this figure with the experimentally observed value of 9×10^{-9}, we see that there is fair agreement.

This relaxation mechanism predicts that T_1 should be proportional to T for temperatures above the Debye temperature and $1/T^3$ below the Debye temperature. The resistivity varies at $1/T^5$ but the extra factor of $1/T^2$ discriminates against forward collisions, which do not contribute to the resistivity but are equally effective in flipping spins. Sodium has a Debye temperature of 200°K, so that according to Elliott the relaxation time should start to fall off faster than T below this temperature. This is in variance with the experimental results.†

The attractive features of Elliott's mechanism are that it explains the shortening of the relaxation time with increasing spin-orbit coupling, which is in accord with the observed results in the alkali metals. If the spin-orbit coupling is too large, the line would be broadened to such an extent that it would be impossible to observe a resonance. All metals in which we were unsuccessful in observing a resonance had larger spin-orbit couplings than the ones in which a resonance was observed. This mechanism also explains the importance of impurities having a larger spin-orbit coupling than the host metal. The presence of impurities would certainly be most important for Li which has the smallest spin-orbit coupling of all the metals, and thus explains the observed results very well.

D. Electronic g Values

All the g determinations were made at 3 cm by comparison with the organic free radical diphenyl picryl hydrazyl.[8] For Li and Na the free radical was introduced into the cavity simultaneously with the bulk metal sample and covered its entire area. This eliminated any errors arising from field inhomogeneities. Since the g values of Be and the free radical lie very close together, the two resonance lines overlapped and alternate runs, with and without the free radical, had to be made. In each case at least 10 runs were made. The experimental results, together with the theoretically calculated values,[9–11] are listed in Table I. For sodium, Brooks' calculations, based essentially on the atomic spin-orbit coupling constants, agree within the experimental error.

Up to the present time there have been no theoretical calculations made on the g shift in beryllium. In beryllium the $2s$ band may be nearly filled, some electrons spilling over into the next zone. The resonance from those electrons would give a negative g shift, while the holes in the $2s$ band would give a positive g shift. If the electronic transitions into the $2s$ band are fast in comparison to a spin relaxation time, we would not expect to see two separate lines, but a single line as observed. The g of this line would fall between the two values. The fact that in Be the g shift is definitely positive suggests that the energy surfaces at the top of the $2s$ band may be degenerate (or nearly so) in which case we would expect a large positive g shift.[7]

TABLE I. Experimental and theoretical values for the g shift, $\Delta g=(g-2.0023)$.

Metal	Experimental values	Theoretical values
Li	$\Delta g\leq10^{-4}$	-6×10^{-5} (Argyres and Kahn[a])
Na	$\Delta g=-(8\pm2)\times10^{-4}$	-3.7×10^{-4} (Yafet[b])
		-6.6×10^{-4} (Brooks[c])
Be	$\Delta g=+(9\pm1)\times10^{-4}$	No theory

[a] See reference 9.
[b] See reference 10.
[c] See reference 11.

The g shift in Li was too small to observe experimentally. The theoretical value is 10 times smaller than that of Na. If one assumes Elliott's relaxation mechanism, this would mean that the relaxation time in pure Li should be 100 times longer than in sodium, i.e., approximately 10^{-6} sec. The experimentally observed value in the purest Li so far obtained is about 4 times shorter than this. An impurity content of only about 1 part in 10^5 of rubidium or cesium, having spin-orbit couplings roughly 1000 time larger than Li, could account for such a shortening of the relaxation time. However, it should be pointed out that a temperature-dependent relaxation time of 10^{-6} sec should already have shown up in the experimental data on the purest Li (see Fig. 14). The failure to observe such a temperature dependent contribution means that the estimated spin-orbit coupling in Li is too large. A further purification of the Li would be very instructive.‡

E. Absolute Values of the Paramagnetic Susceptibility

In our experiments we obtain the paramagnetic contribution to the susceptibility by comparing the integrated resonance absorption in the metal to that of a known amount of the organic free radical diphenyl picryl hydrazyl. The alkali metal sample is in the form of a film, thin in comparison to a skin depth (1–2 microns), which is evaporated on a sheet of mica. (See Sec. III.) The exact amount of metal is obtained by

† *Note added in proof.*—Y. Yafet (private communication) re-examined Elliotts' mechanism and concluded that T_1 is proportional to $1/T$ also below the Debye temperature. This is essentially due to the overwhelming contribution of low-energy phonons to the relaxation mechanism.

[8] C. A. Hutchison and R. C. Pastor, Phys. Rev. **81**, 1282 (1951).
[9] P. Argyres and A. Kahn (to be published).
[10] Y. Yafet, Phys. Rev. **85**, 478 (1952).
[11] H. Brooks, Phys. Rev. **94**, 1411(A) (1954).

‡ *Note added in proof.*—Recently the g-value of cesium was found to be 1.93 ± 0.02 (G. Feher and R. A. Levy; Berkeley Meeting, December 28, 1954) which compares favorably with Brooks' result of 1.94.

allowing it to react with ethyl alcohol and measuring the amount of hydrogen evolved. However, this work is only in its early beginnings and at present our experimental error exceeds that of other workers.[12] More work is being done at present to improve this method.

In principle one could use the results on thick plates to determine the susceptibility with the aid of Eq. (3.3). However, in this case, surface irregularities and the lack of knowledge of the resistivity of the material will severely limit the accuracy with which the susceptibility can be determined. In spite of this inaccuracy we used beryllium particles (not having thin films available) large in comparison to a skin depth in order to get at least a rough estimate for the susceptibility. The value obtained was $\chi_0 \cong 2\times10^{-7}$ cgs volume units. In this case there entered an additional error due to the fact that the exact particle-size distribution was not known. The only justification to quote such a rough estimate is that to our knowledge there is no other published experimental value for the paramagnetic susceptibility of beryllium.

F. Metals in which no Resonance was Observed

Aluminum and magnesium was investigated at 300 Mc/sec in the temperature range between 4°K and 300°K, palladium and tungsten between 77°K and 300°K. In none of the metals a resonance was observed. From the knowledge of the sensitivity of the equipment, this result establishes an upper limit to the spin relaxation time at room temperature of the above metals of $T_2<5\times10^{-10}$ sec. It is difficult to estimate an upper limit of T_2 at 4°K, since the residual resistance, which strongly affects the diffusion time and hence the expected signal, was not known.

The short relaxation times can be explained easily if we assume Elliott's relaxation mechanism.[7] The spin-orbit couplings in the above metals are at least an order larger than in sodium, which would make the relaxation time approximately 2 orders shorter, i.e., 5×10^{-11} sec. The Russian workers, Al'tshuler, Kurenev, and Salikhov[13] report for tungsten a line width of 40 oersteds ($T_2\approx10^{-9}$ sec) at 240 Mc/sec. In a later article Salikhov's[14] value for tungsten is approximately 1000 oersteds ($T_1\approx5\times10^{-11}$ sec). (No explanation or mention of the earlier value is made.) In his article resonances in 19 metals are reported, all of them having roughly the same line width (500–1000 oersteds). Comparing their results with ours on the alkali metals, there exists an alarming discrepancy. Their line width in lithium is roughly 3 orders broader than ours. Their g value for lithium is 2.92, which makes their g shift at least 2–3 orders longer than ours. Their curve for aluminum shows no trace of a "Dysonian character," which makes it questionable whether they observe resonances due to conduction electrons. It should be mentioned that in our early experiments large resonance signals were observed in aluminum. However, they were traced to ferromagnetic impurities (at the low fields used there is an appreciable hysteresis associated with the signal), and they disappeared when high purity materials were used. Also, care has to be taken in the choice of sample holder. Both glass and polystyrene give resonances at $g=4$ and $g=2$. The signal from the latter is very weak and was only observed at 4°K when an excessive quantity of it was used.

[12] Schumacher, Carver, and Slichter, Phys. Rev. **95**, 1089 (1954).
[13] Al'tshuler, Kurenev, and Salikhov, Doklady Akad. Nauk. S.S.S.R. **84**, 677 (1952).
[14] S. G. Salikhov, Akad. Nauk. S.S.S.R. **16**, No. 6, 748 (1952).

APPENDIX A. JUSTIFICATION FOR NEGLECTING SURFACE RELAXATIONS IN THICK PLATES

Dyson finds for the total spin relaxation time T_2 in particles small in comparison to a skin depth (D-105):

$$1/T_2=1/T_1+1/T_S, \tag{A.1}$$

where T_1 is the usual spin-lattice relaxation time and T_S for small particles is given by (D-108):

$$T_S=(4/3)(r/\epsilon v), \tag{A.2}$$

where r=radius of particles, ϵ=probability of a spin disorientation during a surface collision, and v=velocity of electrons at the top of the Fermi surface. For our case (see Sec. IV),

$$T_S=2\times10^{-8}\text{ sec},\quad r\cong2\times10^{-4}\text{ cm},\quad v=10^8\text{ cm sec}^{-1},$$

which gives a value for $\epsilon\cong10^{-4}$.

The important parameter describing the effect of surface relaxations on the line shape is ψ. (D-110)

$$\psi=\epsilon[\tfrac{3}{8}T_2/\tau]^{\frac{1}{2}}, \tag{A.3}$$

where τ=relaxation time obtained from resistivity measurements. For ψ small in comparison to 1, the change in line shape due to surface relaxations may be neglected. Substituting the experimental values into (A.3) we get $\psi=0.03$.

ACKNOWLEDGMENTS

We would like to express our thanks to Professor C. Kittel, for several helpful suggestions. We gratefully acknowledge the cooperation of the Aluminum Corporation of America and the Dow Chemical Company, who prepared for us the high-purity aluminum and magnesium foils. We wish to thank M. Browne for his general assistance, especially the evaporation work; Dr. J. Ubbink for his assistance on many low-temperature runs; also Dr. T. W. Griswold who collaborated in the early experiments; and M. Levy who computed most of the Dyson curves. This work was assisted in part by the Office of Naval Research and the U. S. Signal Corps.

Electron Spin Resonance Absorption in Metals. II. Theory of Electron Diffusion and the Skin Effect*

FREEMAN J. DYSON†
Department of Physics, University of California, Berkeley, California
(Received December 2, 1954)

The theory of paramagnetic resonance absorption by the conduction electrons in a metal is worked out, taking into account the diffusion of the electrons in and out of the thin skin into which the radio-frequency field penetrates. Calculations are carried through in detail for the case of a flat metal plate. It is found that the diffusion has no marked effect on the width of the resonance absorption line, but has a radical effect on the shape of the line. In particular, for a piece of metal thick compared to the skin depth and with a relaxation time long compared to the diffusion time, the line is antisymmetrical about its center and has an unusual characteristic shape.

I. INTRODUCTION

THE solid-state physics group at Berkeley has observed paramagnetic resonances arising from the conduction electrons in metals. Their experimental findings are reported in the preceding paper.[1] It was realized from the beginning of the conduction electron resonance work[2] that the diffusion of electrons in and out of the skin of the metal would have a decisive effect on the shape and intensity of the observed resonance lines. The present paper contains a quantitative theory of this effect, which the author worked out while attached to the Berkeley group during the summer of 1953. The comparison of theory with experimental results is not made here as it is included in the experimental paper.[1]

We build the theory upon a highly idealized model of the electrons in the metal. The electrons are assumed to diffuse like free particles, and the electron magnetic moments are treated like free-particle moments. This is necessary in order to make the theory of the diffusion effect reasonably simple. The comparison of the theory with experiment will then show to what extent the idealized picture is correct. The interest of the whole analysis will ultimately lie, not in the diffusion effect itself, but in the information which the experiments will give concerning the nature of the metallic state after the diffusion effect has been correctly taken into account.

The following is the plan of the present paper. In Sec. II the theoretical model is exactly described. Section III gives a sketch of the physical processes which we believe determine the line width; it is hoped that this rough discussion will make the subsequent mathematics easier to follow. Section IV begins the quantitative treatment. In Sec. V we give a simple theory which is mathematically exact for the limiting case of thick samples and long relaxation times, and which is valid whether the skin effect is in the classical or the anomalous range. This simple theory already contains the most important results of the whole investigation, and will be sufficient for many practical applications. Sections VI–VIII contain a more elaborate theory which is exact for a flat metal plate of any thickness and for any value of the relaxation time, but restricted to the domain of classical skin effect. The extension of this theory to anomalous skin-effect conditions would be possible in principle, but it is not done here. Section IX summarizes the results of the whole paper.

II. THE THEORETICAL MODEL

We suppose a piece of metal of volume V to be placed in a resonant cavity in which there is a radio-frequency magnetic field given by

$$\mathbf{H}_1(\mathbf{r})e^{-i\omega t}+\text{c.c.} \tag{1}$$

Here c.c. stands for "complex conjugate," ω is a fixed frequency, and $\mathbf{H}_1(\mathbf{r})$ is a complex vector function of the position $\mathbf{r}$. The whole cavity lies in a uniform magnetic field $\mathbf{H}$, which is varied slowly during the course of an experiment. The resonance frequency between the two spin states of an electron in the field $\mathbf{H}$ is

$$\nu=2\mu|\mathbf{H}|/\hbar, \tag{2}$$

where μ is the effective electron magnetic moment (not necessarily equal to the value for a free electron). We write

$$\alpha=\omega-\nu, \quad \gamma=\omega+\nu. \tag{3}$$

Under the influence of the field (1), a certain macroscopic magnetization,

$$\mathbf{M}(\mathbf{r})e^{-i\omega t}+\text{c.c.}, \tag{4}$$

will be created in the metal as a result of the turning of the magnetic moments of the conduction electrons. The field $\mathbf{H}_1(\mathbf{r})$ in turn depends on $\mathbf{M}$, because the penetration of radio-frequency field into the metal is changed by the magnetization. It is to be expected that $\mathbf{M}$ will show a resonant behavior and will become large

* Research supported by the Office of Naval Research and the U. S. Signal Corps.
† Now at the Institute for Advanced Study, Princeton, New Jersey.
[1] G. Feher, and A. F. Kip, preceding paper [Phys. Rev. 97, 337 (1955)]. The notations in the paper by Feher and Kip and in the present paper are somewhat different; their ω_0, χ_0, T_D, T_2 are respectively equal to our ν, χ, D, U; other symbols are the same in the two papers.
[2] Griswold, Kip, and Kittel, Phys. Rev. 88, 951 (1952); Kip, Griswold, and Portis, Phys. Rev. 92, 544 (1953).

when ν is nearly equal to ω; in the neighborhood of the resonance the distribution of $\mathbf{H}_1$ will also change. The experimentally observed quantity is the change in the Q-value of the cavity as ν or $\mathbf{H}$ varies. This quantity is a measure of the total energy absorbed in the metal both by eddy currents and by the resistive out-of-phase component of the magnetization. To interpret the experiment, we therefore need to calculate both $\mathbf{H}_1$ and $\mathbf{M}$ as functions of $\mathbf{H}$. The calculation naturally divides itself into two stages: first, to calculate the magnetization produced by a given $\mathbf{H}_1(\mathbf{r})$; second, to find self-consistent solutions of Maxwell's equations with this relation between $\mathbf{H}_1$ and $\mathbf{M}$. The problem is unusual because the magnetization is carried around by the electrons as they diffuse in the metal. Thus $\mathbf{M}$ at a given place and time depends on $\mathbf{H}_1$ at neighboring points and at earlier times.

The electrons which carry the magnetization are assumed all to lie at the top of the Fermi distribution of the conduction electrons and to move with constant velocity v. Let $P(\phi)$ be the probability per unit distance travelled and per unit solid angle that an electron changes its direction by an angle ϕ as a result of collision with lattice vibrations, other electrons, etc. Then each electron "loses its memory" of its direction of motion in a mean free path defined by

$$\Lambda = P^{-1}, \tag{5}$$

where P is the "transport collision probability,"

$$P = \int P(\phi)(1-\cos\phi)d\Omega. \tag{6}$$

A mean collision time τ is defined by

$$\tau = \Lambda/v. \tag{7}$$

We assume each electron to move as an independent classical particle, the changes of direction being random. If $F(\mathbf{r},t)$ is the probability distribution for the position $\mathbf{r}$ of an electron at time t, then over times large compared with τ the classical diffusion equation,

$$\partial F/\partial t = \tfrac{1}{3}v\Lambda\Delta F, \tag{8}$$

is satisfied. The boundary condition is

$$\mathbf{n}\cdot\mathrm{grad}F = 0, \tag{9}$$

where $\mathbf{n}$ is the normal vector at any point on the metal surface.

It is supposed that the spin of each electron is an independent quantum variable, which is only very weakly coupled to the electron's orbital motion. The spin is in general unaffected by collisions and reacts only to the local magnetic field at the place where the electron happens to be. However, the weak spin-orbit coupling exists and very occasionally, about once in every 10^5 collisions in the case of sodium, causes the spin state to change during a collision.[3] To take this effect into account, we do not introduce the spin-orbit coupling explicitly, but we assume a relaxation-time U with the following property. Given any time interval of length t, there is a probability $\exp(-t/U)$ that the spin state of an electron will be undisturbed by collisions during that interval, and a probability $[1-\exp(-t/U)]$ that the final spin state will be randomly distributed relative to the initial spin state.

The existence of such a relaxation time U produces a finite line width proportional to U^{-1} in the resonance signal from nondiffusing electrons. This we call the "natural width" of the line. Any other effects besides spin-orbit coupling which contribute to the natural width are supposed also included in U.

In Sec. VIII a more general model is introduced, in which the electrons have a finite probability of spin disorientation when they strike the metal surface, in addition to the steady relaxation effects included in U. For simplicity, the calculations are carried through in Secs. IV–VII for the model without surface relaxation. Exact results including surface relaxation are to be found at the end of Sec. VIII.

It is assumed that the electron magnetization is always far from saturation, so that $\mathbf{M}$ is small compared with $\mathbf{H}_1$. Terms of second and higher order in the static magnetic susceptibility χ of the metal are therefore neglected throughout the calculation. This makes it unnecessary to discuss the difficult question, whether the radio-frequency field which effectively acts on the electron spins is the field $\mathbf{H}_1$, or the magnetic induction $\mathbf{B}$, or something intermediate.[4] With an error at most of the order χ^2, we suppose $\mathbf{H}_1$ to be the effective field acting on the spins.

The penetration of the radio-frequency field into the metal will be limited by skin effect. The classical skin depth is given by

$$\delta = (c^2/2\pi\sigma\omega)^{\frac{1}{2}}, \tag{10}$$

where σ is the conductivity of the metal. If δ is large compared to Λ, we are in the domain of "classical skin effect" and the field penetration is controlled by δ. If δ is not large compared to Λ, we have "anomalous skin effect"[5] and the penetration is controlled by Λ.

III. QUALITATIVE ARGUMENT

It is convenient in what follows to use the symbol δ to mean the actual skin depth, given by (10) in the classical case and equal to Λ in the anomalous case. According to Eq. (8), an average electron will diffuse across the skin depth in a time of the order of

$$D = (3\delta^2/2v\Lambda). \tag{11}$$

[3] This mechanism for the spin relaxation was proposed by R. J. Elliott, Phys. Rev. **96**, 266, 280 (1954). Reasons for believing that it is the controlling mechanism are given in the paper by G. Feher and A. F. Kip, reference 1, Sec. IV. C.2.

[4] For a discussion of this point see G. H. Wannier, Phys. Rev. **72**, 304 (1947).

[5] G. Reuter and E. Sondheimer, Proc. Roy. Soc. (London) **A195**, 336 (1948).

The amplitude and phase of the field $\mathbf{H}_1$ will change by factors of the order of e within the distance δ. Therefore the average electron will feel the radio-frequency field not with a fixed frequency ω but with frequencies spread over a band of width D^{-1}. At first glance this argument leads one to believe that diffusion should increase the width of the resonance line from U^{-1} to D^{-1} if $D<U$. The surprising fact is that the observed width remains equal to U^{-1}, although the intensity and shape of the line are changed by the diffusion. The aim of the following discussion is to explain this result in a qualitative way, and to draw an analogy to a more familiar physical example where a similar situation arises.

Suppose that the relaxation time U is long compared with D. The radio-frequency field seen by one electron will have the form

$$F(t)=f(t)e^{-i\omega t}+\text{c.c.}, \tag{12}$$

where $f(t)$ is a modulation factor with a frequency band of the order of D^{-1}. However, $f(t)$ is not a completely random function. It is composed of short pulses of duration D, extending over the time U and separated by irregular intervals. The pulses occur when the electron comes close to the surface and the amplitude of $f(t)$ is large; the intervals occur when the electron is far from the surface and $f(t)$ is almost zero. Two facts are here of decisive importance. (i) Each electron which leaves the surface has a high probability of returning, and so the $f(t)$ for an average electron contains more than one pulse. (ii) The phase of $f(t)$ returns to the same value each time the electron reaches the surface, and so the integral of $f(t)$ over a pulse has a finite average value which is not zero.

Now consider the form of the spectrum of $F(t)$. Each single pulse will give a smooth spectrum spread over a band of width D^{-1}. Two pulses separated by an interval I will give an interference pattern, with maxima and minima separated by a spacing of order I^{-1}, extending over the band width D^{-1}. There will always be constructive interference at the central frequency ω, because of point (ii). Three or more pulses will give a more complicated interference pattern over the band D^{-1}, still with the central maximum at ω. Now let this spectrum be squared to give the radio-frequency power spectrum seen by one electron, and let it then be averaged over many electrons. The interval-length I will vary at random over values from zero up to U. Therefore the interference patterns will average out to a smooth curve of width D^{-1}, except for the central maxima which will add constructively to give a single peak with width of the order of U^{-1}.

Thus we expect the observed absorption spectrum of radio-frequency energy, which is an average over many electrons of the power spectrum acting upon one electron, to have a structure of width U^{-1} superimposed upon a weaker background of width D^{-1}. This is in fact what is observed, although the conditions of the experiment are such as to make only the central structure clearly distinguishable.

The device invented by Ramsey[6] for observing molecular beam resonances with two separated oscillating fields is based on exactly the same principle. In Ramsey's apparatus the molecules pass through two short pulses of radio-frequency field. The phase relation of the two pulses is fixed, but the time interval between them depends on the velocity and so varies from molecule to molecule. The observed resonance line consists of a broad background with width determined by the duration of each pulse, and a narrow central structure with width determined by the interval between them. For obtaining sharp lines, it is essential that the molecular velocity is distributed at random over a wide range. This random element in Ramsey's apparatus plays precisely the same role as the random diffusion of the electrons in Feher's experiment.

The qualitative theory of this section would be exact if the only cause of radio-frequency energy absorption in a metal were the flipping of the electron spins. In fact, however, the eddy-current losses play an equally important part, and these have no analog in Ramsey's experiment. Therefore the qualitative theory explains only the fact that narrow lines are seen, and does not predict the shape of the lines correctly. The shape depends in a complicated way upon the phase relationships of electric and magnetic fields in the metal skin, and a simple physical picture of these effects is still lacking.

IV. CALCULATION OF THE MAGNETIZATION

We now begin the quantitative treatment of the model described in Sec. II. In this section we calculate the magnetization produced by a given magnetic field (1). Consider a single electron whose spin is described by the wave function

$$u(t)=a_1(t)\exp(-\tfrac{1}{2}i\nu t)u_1+a_2(t)\exp(\tfrac{1}{2}i\nu t)u_2. \tag{13}$$

Here u_1, u_2 are the states with spin parallel and antiparallel to the direction of $\mathbf{H}$. The amplitudes a_1, a_2 vary with time by virtue of the interaction energy,

$$\mu\boldsymbol{\sigma}\cdot\mathbf{H}_1(\mathbf{r}(t))e^{-i\omega t}+\text{c.c.}, \tag{14}$$

of the electron spin $\boldsymbol{\sigma}$ with the field (1). The Schrödinger equation for a_1 is

$$i\hbar\dot{a}_1=\mu(\mathbf{H}_1\cdot\mathbf{k}e^{-i\omega t}+\mathbf{H}_1^*\cdot\mathbf{k}e^{i\omega t})a_1 +\mu(\mathbf{H}_1\cdot\mathbf{s}e^{-i\omega t}+\mathbf{H}_1^*\cdot\mathbf{s}e^{i\gamma t})a_2. \tag{15}$$

Here $\mathbf{H}_1$ stands for $\mathbf{H}_1(\mathbf{r}(t))$, $\mathbf{k}$ is a unit vector in the direction of $\mathbf{H}$, and $\mathbf{s}$ is the vector

$$\mathbf{s}=(u_1^*\boldsymbol{\sigma}u_2). \tag{16}$$

Since we are making the approximation of treating the magnetization as linear in $\mathbf{H}_1$, Eq. (15) may be solved by first-order perturbation theory. Consider an

[6] N. F. Ramsey, Phys. Rev. 78, 695 (1950).

electron known to be in the state u_1 at any time t'. At a later time t its spin state will be determined by Eq. (15), and a simple calculation shows that the expectation value of its spin vector will be

$$\bar{\sigma}(t,t')=\mathbf{k}+(i\mu/\hbar)\int_{t'}^{t}du[\mathbf{s}\cdot\mathbf{H}_1(\mathbf{r}(u))e^{-i\alpha u-i\nu t}\mathbf{s}^*-\mathbf{s}^*\cdot\mathbf{H}_1(\mathbf{r}(u))e^{-i\gamma u+i\nu t}\mathbf{s}]+\text{c.c.} \quad (17)$$

For an electron in the state u_2 at time t', the expectation value of the spin at time t is given by the negative of (17). According to our model, there is a probability $U^{-1}\exp[-(t-t')/U]dt'$ that an electron observed at time t suffered its last random change of spin state during the time interval $(t', t'+dt')$. Further, each random change of spin state leads to the states u_1 and u_2 with unequal probabilities p_1 and p_2, because of the thermodynamical preference for the lower-energy state. Hence the weighted average of the spin-expectation value of an electron observed at time t is

$$\begin{aligned}\bar{\sigma}(t)&=(p_1-p_2)U^{-1}\int_{-\infty}^{t}\exp[-(t-t')/U]\bar{\sigma}(t,t')dt'\\&=(p_1-p_2)\mathbf{k}+(i\mu/\hbar)(p_1-p_2)\\&\quad\times\int_{-\infty}^{t}\exp[-(t-u)/U]du[\mathbf{s}\cdot\mathbf{H}_1(\mathbf{r}(u))\\&\quad\times e^{-i\alpha u-i\nu t}\mathbf{s}^*-\mathbf{s}^*\cdot\mathbf{H}_1(\mathbf{r}(u))e^{-i\gamma u+i\nu t}\mathbf{s}]+\text{c.c.} \quad (18)\end{aligned}$$

We wish to calculate the macroscopic intensity of magnetization produced by the electrons observed in the neighborhood of a point $\mathbf{r}$ at time t. This is obtained if we average Eq. (18) over the possible previous histories of an electron which happens to be at $\mathbf{r}$ at time t. So let $G(\mathbf{r}',u,\mathbf{r},t)$ be the probability that an electron observed at $\mathbf{r}$ at time t was at $\mathbf{r}'$ at time u. Let N be the number of electrons per unit volume taking part in the magnetization. The product $N(p_1-p_2)$ is directly related to the static susceptibility χ of the metal by

$$\chi|\mathbf{H}|=\tfrac{1}{2}\chi\nu\hbar/\mu=-N\mu(p_1-p_2). \quad (19)$$

Averaging (18) over the electron histories and using (19), the magnetization density becomes

$$\begin{aligned}\mathbf{M}(\mathbf{r},t)&=\chi|\mathbf{H}|\mathbf{k}+\tfrac{1}{2}i\chi\nu\int_{-\infty}^{t}\exp[-(t-u)/U]du\\&\quad\times\int_V G(\mathbf{r}',u,\mathbf{r},t)d\mathbf{r}'[\mathbf{s}\cdot\mathbf{H}_1(\mathbf{r}')e^{-i\alpha u-i\nu t}\mathbf{s}^*\\&\quad-\mathbf{s}^*\cdot\mathbf{H}_1(\mathbf{r}')e^{-i\gamma u+i\nu t}\mathbf{s}]+\text{c.c.} \quad (20)\end{aligned}$$

This is the required nonlocal relation between $\mathbf{M}$ and $\mathbf{H}_1$.

It is required now to find solutions of the Maxwell equations consistent with (20). Assuming solutions varying with time like $\exp(-i\omega t)$ as in Eqs. (1) and (4), the Maxwell equations are

$$\operatorname{curl}\mathbf{H}_1=(4\pi/c)\mathbf{j}, \quad (21)$$

$$\operatorname{curl}\mathbf{E}=-(1/c)(d\mathbf{B}/dt)=(i\omega/c)\mathbf{B}, \quad (22)$$

$$\mathbf{B}=\mathbf{H}_1+4\pi\mathbf{M}. \quad (23)$$

Here $\mathbf{M}$ is given by Eq. (20) inside the metal and $\mathbf{M}=0$ outside. The current $\mathbf{j}$ is

$$\mathbf{j}=\sigma\mathbf{E}, \quad (24)$$

if the spin effect is in the classical range. When the spin effect is anomalous, $\mathbf{j}$ is a more complicated function[5] of $\mathbf{E}$ but is still independent of $\mathbf{M}$. The displacement current is in any case negligible inside the metal.

The experimentally observed quantity is the radio-frequency energy P absorbed into the metal per unit time and per unit area of surface. This is the real part of the normal component of the complex Poynting vector taken just outside the surface. It is convenient to define the complex surface impedance,

$$Z=(4\pi/c)[\mathbf{n}\cdot(\mathbf{E}_0\times\mathbf{H}_{10}^*)]/|\mathbf{H}_{10}|^2, \quad (25)$$

where $\mathbf{E}_0$ and $\mathbf{H}_{10}$ are the electric and magnetic fields immediately outside the surface, and $\mathbf{n}$ is the unit inward normal. This quantity Z is characteristic of the metal surface and depends on the direction of polarization but not on the magnitude of the field $\mathbf{H}_1$. The absorption rate P is then given by

$$P=(c/4\pi)^2|\mathbf{H}_{10}|^2(\operatorname{Re}Z). \quad (26)$$

The imaginary part of Z gives the corresponding inductive load on the cavity, and would in principle be observable by a small shift in the resonant frequency of the cavity as the steady field $\mathbf{H}$ is varied.

V. SIMPLE THEORY FOR THICK SAMPLES AND LONG RELAXATION TIMES

In the case of a metal sample which has dimensions large compared to the skin depth (either classical or anomalous), we may idealize the problem by considering the surface to be the plane $z=0$, the metal occupying the half-space $z>0$, and the field $\mathbf{H}_1$ to be a function of z only. The integration over the x and y coordinates in Eq. (20) is then trivial. We may replace G by the Green's function for the diffusion equations (8) and (9) in one dimension, namely

$$\begin{aligned}G(z',u,z,t)&=[(4/3)\pi v\Lambda|t-u|]^{-\frac{1}{2}}\\&\quad\times\{\exp(-3(z-z')^2/4v\Lambda|t-u|)\\&\quad+\exp(-3(z+z')^2/4v\Lambda|t-u|)\}. \quad (27)\end{aligned}$$

The integration over $\mathbf{r}'$ in Eq. (20) is replaced by an integration over z' from 0 to ∞.

In addition we make the approximation of a long relaxation time, that is to say we assume U to be long compared with the diffusion time D given by (11). This means that the time integration in Eq. (20) extends effectively over a time long compared with D, while the

space integration extends only over values of z and z' of the order of δ. Comparing Eq. (27) with (11), the exponents in Eq. (27) will be small over the major part of the space and time integrations of Eq. (20). Therefore in the limit of long relaxation time we may replace both exponentials in Eq. (27) by 1 simply. The physical meaning of this approximation is clear; for long relaxation times the magnetization produced by the radio-frequency field at a depth z' in the skin will diffuse evenly over the skin in a way which becomes eventually independent of z'. Another consequence of the long relaxation time approximation is that the nonresonant term in γ in Eq. (20) becomes negligible compared with the resonant term in α. Substituting Eq. (27) into (20) with these approximations, the radio-frequency magnetization is equal to (4), with $\mathbf{M}(\mathbf{r})$ given by

$$\mathbf{M}(\mathbf{r})=\tfrac{1}{2}i\chi\nu\int_0^\infty dw[\tfrac{1}{3}\pi v\Lambda w]^{-\frac{1}{2}}\exp[-w(U^{-1}-i\alpha)]\times\int_0^\infty dz'\mathbf{s}\cdot\mathbf{H}_1(z')\mathbf{s}^*, \quad (28)$$

independent of the position $\mathbf{r}$. The time and space integrations are now separated and can be easily performed.

With error of order χ^2, we replace $\mathbf{H}_1$ by $\mathbf{B}$ in Eq. (28). The second Maxwell equation (22) becomes, for fields depending only on z,

$$(i\omega/c)\mathbf{B}=(d/dz)(\mathbf{n}\times\mathbf{E}), \quad (29)$$

and therefore

$$\int_0^\infty \mathbf{B}(z')dz'=(ic/\omega)(\mathbf{n}\times\mathbf{E}_0). \quad (30)$$

Thus Eq. (28) reduces to

$$\mathbf{M}(\mathbf{r})=-K(1-ix)^{-\frac{1}{2}}[\mathbf{s}\cdot(\mathbf{n}\times\mathbf{E}_0)]\mathbf{s}^*, \quad (31)$$

with

$$K=(\chi\nu c/2\omega)(3U/v\Lambda)^{\frac{1}{2}}, \quad (32)$$

$$x=\alpha U=(\omega-\nu)U, \quad (33)$$

and the complex square root in Eq. (31) to be taken with positive real part so that we have explicitly:

$$[1-ix]^{-\frac{1}{2}}=2^{-\frac{1}{2}}(1+x^2)^{-\frac{1}{4}}[\eta+i\xi], \quad (34)$$

$$\eta=[(1+x^2)^{\frac{1}{2}}+1]^{\frac{1}{2}}, \quad \xi=(\mathrm{sgn}x)[(1+x^2)^{\frac{1}{2}}-1]^{\frac{1}{2}}. \quad (35)$$

By making use[7] of known results on skin effect, we can now complete the solution of the Maxwell equations with the constant magnetization $\mathbf{M}(\mathbf{r})$. The Eqs. (21) and (22), when written in terms of $\mathbf{B}$ and $\mathbf{E}$, are completely independent of the magnetization. This is true both with normal and with anomalous skin effect. Therefore the $\mathbf{B}$ and $\mathbf{E}$ fields inside the metal are given precisely by the usual theory of the skin effect without magnetization. The boundary condition at the surface is that the tangential components of $\mathbf{E}$ and $\mathbf{H}_1$ are continuous. Thus the fields immediately outside the surface are

$$\mathbf{E}_0=\mathbf{E}_0, \quad \mathbf{H}_{10}=\mathbf{B}_0-4\pi\mathbf{M}_\perp, \quad (36)$$

where $\mathbf{E}_0$ and $\mathbf{B}_0$ are the fields calculated without magnetization, and $\mathbf{M}_\perp$ is the tangential component of the magnetization given by Eq. (31).

We write Z_0 for the surface impedance of the metal without magnetization, given by

$$Z_0=(4\pi/c)[\mathbf{n}\cdot(\mathbf{E}_0\times\mathbf{B}_0^*)]/|\mathbf{B}_0|^2. \quad (37)$$

Let $\mathbf{b}$ be the complex polarization vector of $\mathbf{B}_0$,

$$\mathbf{b}=\mathbf{B}_0/|\mathbf{B}_0|, \quad (\mathbf{b}\cdot\mathbf{b}^*)=1. \quad (38)$$

By Eqs. (30) and (37), we have

$$(\mathbf{n}\times\mathbf{E}_0)=(c/4\pi)Z_0|\mathbf{B}_0|\mathbf{b}. \quad (39)$$

Substituting (31), (36), and (39) into Eq. (25), the surface impedance with magnetization becomes

$$Z=Z_0[1+Kc(1-ix)^{-\frac{1}{2}}fZ_0]^{-1}=Z_0-Kcf(1-ix)^{-\frac{1}{2}}Z_0^2, \quad (40)$$

neglecting terms of order χ^2. We have written

$$f=|(\mathbf{s}\cdot\mathbf{b})|^2=|(u_1^*\mathbf{b}\cdot\boldsymbol{\sigma}u_2)|^2=(\mathbf{b}\times\mathbf{k})\cdot(\mathbf{b}^*\times\mathbf{k})+i(\mathbf{b}\times\mathbf{b}^*)\cdot\mathbf{k}, \quad (41)$$

where $\mathbf{k}$ is the unit vector parallel to the steady field $\mathbf{H}$. In general, f is a real number lying between 0 and 2, measuring the proportion of the radio-frequency field which is circularly polarized in the positive sense about the direction of $\mathbf{H}$. The maximum value 2 is attained by f when $\mathbf{b}=2^{-\frac{1}{2}}(1,i,0)$, i.e., for pure circularly polarized radiation. In practice we shall usually be dealing with linearly polarized fields, in which case $\mathbf{b}$ is real and

$$f=\sin^2\phi, \quad (42)$$

where ϕ is the angle between the polarization directions of $\mathbf{H}_1$ and $\mathbf{H}$.

When the skin effect is anomalous, the Reuter-Sondheimer theory[5] will give the value of Z_0 to be used in Eq. (40). When the skin effect is classical, Eqs. (21), (22), and (24) give the classical result

$$Z_0=(1-i)(\sigma\delta)^{-1}. \quad (43)$$

Substituting this into (40), using Eqs. (26), (32), (10), and (11), we have the result

$$P=\tfrac{1}{4}\delta\omega|\mathbf{H}_{10}|^2[(1/2\pi)-\chi\nu f(DU)^{\frac{1}{2}}\xi(1+x^2)^{-\frac{1}{4}}], \quad (44)$$

with x and ξ given by Eq. (33) and (35). Thus the absorption line, in the case of a thick sample with long relaxation time and classical skin effect, has the unusual and characteristic shape given by the function $\xi(1+x^2)^{-\frac{1}{4}}$. Since ξ is an odd function of $(\omega-\nu)$, the line is in this approximation antisymmetrical about its center. The width is proportional to U^{-1} and independent of the diffusion time D. These are the most

[7] The subsequent argument is due to Professor Kittel (private communication).

striking conclusions of the theory. The experiments[1] support them qualitatively; however for a quantitative comparison with experiment the discussion of a limiting case is insufficient, and the more detailed theory of the following sections is required.

VI. SOLUTION OF MAXWELL EQUATIONS FOR A FLAT PLATE

To give an exact theory for a metal sample which is not necessarily large or small compared with the skin depth, we must choose a definite shape for the metal. We shall confine ourselves to the case of a flat plate, which is a convenient shape both experimentally and theoretically. Other shapes could be handled by similar methods. We make no restriction on the relative magnitude of D and U, but we shall assume henceforth that we are in the domain of classical skin effect, i.e., that Eq. (24) is valid.

The metal will again be considered to be infinite in the x and y directions, thus it occupies the volume $0\leq z\leq\theta$ where θ is the thickness. An important parameter in the problem will be the ratio of thickness to skin depth,

$$\lambda=\theta/\delta. \tag{45}$$

It is also convenient to introduce the dimensionless ratios:

$$R=(D/U)^{\frac{1}{2}}, \tag{46}$$

$$a=\lambda R=(3\theta^2/2v\Delta U)^{\frac{1}{2}}. \tag{47}$$

As before, we assume the fields to depend on z only, so that the problem is essentially one-dimensional. We shall also make the arbitrary assumption that the magnetic field is equal on the two sides of the plate, i.e., an even function of $(z-\frac{1}{2}\theta)$. This assumption is made only to simplify the algebra; an unsymmetric field could be calculated in exactly the same way. In particular, for a thick plate ($\theta\gg\delta$) the fields on the two sides are independent and the symmetric solutions apply unchanged to unsymmetric conditions; for example, the symmetric solutions may be used when the plate is thick and forms part of the wall of the cavity.

The idea of the following analysis is to reduce Eq. (20) to a tractable form by introducing the eigenfunctions $\psi_n(\mathbf{r})$, $n=0, 1, 2, \cdots$; these are defined as a complete orthogonal set of functions satisfying the eigenvalue equation

$$\Delta\psi_n=-\mu_n\psi_n, \tag{48}$$

with the boundary condition

$$\mathbf{n}\cdot\operatorname{grad}\psi_n=0, \tag{49}$$

and the normalization

$$\int_V\psi_m^*\psi_n d\mathbf{r}=\delta_{mn}. \tag{50}$$

In particular, the first eigenfunction is a constant:

$$\psi_0(\mathbf{r})=V^{-\frac{1}{2}},\quad \mu_0=0. \tag{51}$$

The Green's function G which appears in Eq. (20) satisfies Eqs. (8) and (9) as a function of its arguments $(\mathbf{r},t)$ and reduces to a δ function, $\delta(\mathbf{r}-\mathbf{r}')$, for $t=u$. These conditions determine G as the series expansion:

$$G=\textstyle\sum_n\psi_n^*(\mathbf{r}')\psi_n(\mathbf{r})\exp[-\frac{1}{3}v\Lambda\mu_n(t-u)]. \tag{52}$$

Let the field $\mathbf{H}_1(\mathbf{r})$ inside the metal also be expanded in a series

$$\mathbf{H}_1(\mathbf{r})=\textstyle\sum_n\mathbf{h}_n\psi_n(\mathbf{r}),\quad \mathbf{h}_n=\int_V\psi_n^*\mathbf{H}_1 d\mathbf{r}, \tag{53}$$

where the coefficients $\mathbf{h}_n$ are vectors. Substituting from Eq. (52), (53) into Eq. (20), the space and time integrations can be performed immediately. The radio-frequency magnetization then appears in the form (4) with

$$\mathbf{M}(\mathbf{r})=\tfrac{1}{2}\nu\chi i\textstyle\sum_n\psi_n(\mathbf{r})[\eta_n(\mathbf{s}\cdot\mathbf{h}_n)\mathbf{s}^*-\zeta_n(\mathbf{s}^*\cdot\mathbf{h}_n)\mathbf{s}], \tag{54}$$

$$\eta_n=[U^{-1}+\tfrac{1}{3}v\Lambda\mu_n-i\alpha]^{-1},\quad \zeta_n=[U^{-1}+\tfrac{1}{3}v\Lambda\mu_n-i\gamma]^{-1}. \tag{55}$$

The nonlocal relation between $\mathbf{M}$ and $\mathbf{H}_1$ given by Eq. (54), (55) is completely general, holding for a metal of any shape and for any kind of skin effect.

Now we specialize to the flat plate with classical skin effect and symmetrical fields which are functions of z only. The eigenfunctions which appear in Eq. (53) are then

$$\psi_n=[\epsilon_n/\theta]^{\frac{1}{2}}\cos[2\pi nz/\theta],\quad \mu_n=4\pi^2n^2/\theta^2, \tag{56}$$

where

$$\epsilon_0=1,\quad \epsilon_n=2 \text{ for } n>0, \tag{57}$$

and the normalization is per unit area of the plate. The Maxwell equations (21), (22), (23), (24) give, inside the metal,

$$\begin{aligned}2i\delta^{-2}(\mathbf{H}_1+4\pi\mathbf{M})&=(4\pi\sigma/c)(d/dz)(\mathbf{n}\times\mathbf{E})\\&=(d^2/dz^2)[\mathbf{n}\times(\mathbf{n}\times\mathbf{H}_1)]=-(d^2/dz^2)\mathbf{H}_\perp,\end{aligned} \tag{58}$$

where $\mathbf{H}_\perp$ is the component of $\mathbf{H}_1$ perpendicular to $\mathbf{n}$. Let $\mathbf{M}_\perp$, the component of $\mathbf{M}$ perpendicular to $\mathbf{n}$, be expressed as a series expansion

$$\mathbf{M}_\perp(z)=(1/4\pi)\textstyle\sum_n\mathbf{m}_n\psi_n(z). \tag{59}$$

We multiply Eq. (58) by $\psi_n^*(z)$ and integrate over z between two limits lying just inside the surface at $z=0$, $z=\theta$. The right side is then integrated by parts twice, using Eqs. (49) and (48). The result is

$$\begin{aligned}&2i\delta^{-2}(\mathbf{h}_{n\perp}+\mathbf{m}_n)\\&=(4\pi\sigma/c)\psi_n^*(0)[(\mathbf{n}\times\mathbf{E}_\theta)-(\mathbf{n}\times\mathbf{E}_0)]\\&\qquad-\int(d^2\psi_n^*/dz^2)\mathbf{H}_\perp dz\\&=-(8\pi\sigma/c)[\epsilon_n/\theta]^{\frac{1}{2}}(\mathbf{n}\times\mathbf{E}_0)+\mu_n\mathbf{h}_{n\perp}.\end{aligned} \tag{60}$$

Here $\mathbf{h}_{n\perp}$ is the component of $\mathbf{h}_n$ perpendicular to $\mathbf{n}$. The electric fields at the surfaces are $\mathbf{E}_0$ and $\mathbf{E}_\theta=-\mathbf{E}_0$ because of the assumed symmetry.

The information we need from the Maxwell equations is all contained in Eq. (60), while Eq. (54) gives the coefficients $\mathbf{m}_n$ in terms of the $\mathbf{h}_n$. One could solve Eq. (60) and (54) for the $\mathbf{h}_n$ and thus obtain the complete field-distribution inside the metal in terms of the electric field $\mathbf{E}_0$ at the surface.

The solution of the equations becomes simple because all the $\mathbf{m}_n$ are proportional to χ and may be treated as small. With an error of order χ, Eq. (60) implies that all the $\mathbf{h}_n$ are parallel to the vector $(\mathbf{n}\times\mathbf{E}_0)$, which is parallel to the polarization vector $\mathbf{b}$ defined by Eq. (38). Therefore we may write in Eq. (54), with an error of order χ^2,

$$(\mathbf{s}\cdot\mathbf{h}_n)=(\mathbf{s}\cdot\mathbf{b})(\mathbf{b}^*\cdot\mathbf{h}_n),\quad (\mathbf{s}^*\cdot\mathbf{h}_n)=(\mathbf{s}^*\cdot\mathbf{b})(\mathbf{b}^*\cdot\mathbf{h}_n). \tag{61}$$

A comparison of Eq. (54) with (59), using Eq. (61), gives the relation

$$(\mathbf{b}^*\cdot\mathbf{m}_n)=g_n(\mathbf{b}^*\cdot\mathbf{h}_n), \tag{62}$$

with g_n defined by

$$g_n=2\pi\nu\chi i[f\eta_n-f'\zeta_n]. \tag{63}$$

Here f is defined by Eq. (41) and f' by

$$f'=|(\mathbf{s}^*\cdot\mathbf{b})|^2=(\mathbf{b}\times\mathbf{k})\cdot(\mathbf{b}^*\times\mathbf{k})-i(\mathbf{b}\times\mathbf{b}^*)\cdot\mathbf{k}, \tag{64}$$

so that for linearly polarized fields

$$f'=f=\sin^2\phi, \tag{65}$$

with ϕ defined as in Eq. (42). Taking the scalar product of Eq. (60) with the vector $\mathbf{b}^*$ which is perpendicular to $\mathfrak{n}$, and using Eq. (62), we obtain

$$(8\pi\sigma/c)[\epsilon_n/\theta]^{\frac{1}{2}}[\mathbf{b}^*\cdot(\mathbf{n}\times\mathbf{E}_0)] =[\mu_n-2i\delta^{-2}(1+g_n)](\mathbf{b}^*\cdot\mathbf{h}_n). \tag{66}$$

The radio-frequency power absorbed per unit area of metal surface is given by Eq. (26) and (25). Since $\mathbf{H}_{10}^*$ is almost parallel to $\mathbf{b}^*$, with an error of order χ^2 the surface impedance Z may be written

$$Z=(4\pi/c)[\mathbf{b}^*\cdot(\mathbf{n}\times\mathbf{E}_0)]/(\mathbf{b}^*\cdot\mathbf{H}_{10}). \tag{67}$$

Since $\mathbf{H}_\perp$ is continuous across the surface, Eqs. (53) and (56) give

$$(\mathbf{b}^*\cdot\mathbf{H}_{10})=\sum_n[\epsilon_n/\theta]^{\frac{1}{2}}(\mathbf{b}^*\cdot\mathbf{h}_n). \tag{68}$$

Substituting from Eqs. (68) and (66) into (67) gives the final explicit formula for Z:

$$Z^{-1}=(2\sigma/\theta)\sum_n\epsilon_n[\mu_n-2i\delta^{-2}(1+g_n)]^{-1}, \tag{69}$$

or, neglecting terms in χ^2,

$$Z^{-1}=(2\sigma/\theta)\sum_n\epsilon_n[(\mu_n-2i\delta^{-2})^{-1} +2i\delta^{-2}g_n(\mu_n-2i\delta^{-2})^{-2}]. \tag{70}$$

VII. EVALUATION OF THE LINE SHAPE FOR FOR A FLAT PLATE

In this section we first calculate in closed form the sum of the series (70). We then discuss in detail the shape of the observed resonance line for the various ranges of values of the parameters which can occur experimentally.

Using Eqs. (56) and (57), the summation in (70) may be replaced by a summation over integers n from minus to plus infinity. Taking the inverse of Eq. (70) and again neglecting terms of order χ^2,

$$Z=(2/\sigma\theta)\{F+\tfrac{1}{2}\pi\nu\chi a^2UF^2[fG(\alpha)-f'G(\gamma)]\}, \tag{71}$$

$$F=[\sum_{-\infty}^{\infty}(\pi^2n^2-\tfrac{1}{2}i\lambda^2)^{-1}]^{-1}, \tag{72}$$

$$G(\alpha)=\lambda^2\sum_{-\infty}^{\infty}(\pi^2n^2-\tfrac{1}{2}i\lambda^2)^{-2}[\pi^2n^2+\tfrac{1}{2}a^2(1-ix)]^{-1}, \tag{73}$$

where x, λ, a are given by Eqs. (33), (45), (47), respectively. Now introducing the notations

$$u=\tfrac{1}{2}\lambda(1+i),\quad u^2=\tfrac{1}{2}i\lambda^2, \tag{74}$$

$$w=\tfrac{1}{2}a(\xi+i\eta),\quad w^2=\tfrac{1}{2}a^2(ix-1), \tag{75}$$

with ξ and η given by Eq. (36), the sums (72) and (73) become

$$F=-u\tan u, \tag{76}$$

$$G(\alpha)=\frac{i}{(w^2-u^2)^2}\left[2u^2\frac{\cot w}{w}+(w^2-3u^2)\frac{\cot u}{u} +(w^2-u^2)\operatorname{cosec}^2u\right]. \tag{77}$$

The real part of Z, determined by Eqs. (71), (76), and (77), gives the explicit form of the line shape for any values of the ratios λ and a. We now proceed to consider the simplifying approximations which can be made for values of λ and a lying in various ranges.

(a) Thin Case

$$\lambda\leq 4,\quad \text{or}\quad \theta\leq 4\delta.$$

For $\lambda\leq 1$, the term $n=0$ is much the greatest in the series (72) and (73), and we may neglect the other terms. Hence Eq. (71) becomes

$$Z=\frac{2\pi\theta\omega}{c^2}\left[-i+2\pi\nu\chi U\left(\frac{f}{1-iU\alpha}-\frac{f'}{1-iU\gamma}\right)\right]. \tag{78}$$

This result is entirely independent of electron diffusion. The energy absorption, which is proportional to the real part of Z, varies with $\mathbf{H}$ approximately as $[1+U^2\alpha^2]^{-1}$, and so gives an absorption line of the ordinary symmetrical Lorentz shape with the natural half-width U^{-1}.

The result (78) will still be reasonably accurate up to $\lambda=4$. The higher terms in (73), being small and spread out over a width greater than U^{-1}, will hardly be observable.

(b) Intermediate Case

$$\lambda \geq 4, \quad a \leq 4.$$

Experimentally the relaxation time U is of the order of 10^{-8} sec, while the diffusion time D is not shorter than 10^{-10} sec. So the ratio $R=(D/U)^{\frac{1}{2}}=(a/\lambda)$ is not less than about 1/10. Therefore, this intermediate case can occur only when $4 \leq \lambda \leq 40$, i.e., when the metal is thick but not very thick compared to the skin depth δ.

When $4 \leq \lambda \leq 10$ it will be convenient to calculate with the first few terms of the series (73), which still converges rapidly although the term $n=0$ is no longer predominant. When $10<\lambda \leq 40$ and $a \leq 4$, it will be more convenient to use Eqs. (76) and (77) with the good approximations

$$\tan u = i, \quad \operatorname{cosec}^2 u = 0. \tag{79}$$

This gives for the central part of the absorption line, where $x=U\alpha \leq 1$, the shape

$$Z=(2/\sigma\theta)[-iu-i\pi\nu\chi a^2 U f(\cot w/w)]. \tag{80}$$

Naturally this intermediate case is more complicated to calculate than either the thin or the thick case. Fortunately, the intermediate case corresponds to metal thicknesses in the range 10^{-4}–10^{-3} cm, which is an awkward range for experimental preparation and will not often occur in practice.

(c) Thick Case

$$\lambda \geq 4, \quad a \geq 4.$$

These conditions are always fulfilled for sufficiently thick samples, say for $\theta > 40\delta$. And the results for this case will be independent of the shape of the metal, so long as the radius of curvature of the surface is everywhere large compared to δ.

The imaginary part of w given by Eqs. (75) and (35) is in this case always large compared to 1. Hence we may put $\cot w = \cot u = -i$ in Eq. (77). This gives for Z the formula

$$Z=(\sigma\delta)^{-1}[1-i-\pi\nu\chi D(fL-f'L')], \tag{81}$$

with

$$L=[R^2(x+i)-1]^{-2}\{4R^{-1}(\xi+i\eta)^{-1} + (1-i)[R^2(x+i)-3]\}, \tag{82}$$

and L' the same as L with α replaced by γ. In the absence of magnetization, Z reduces correctly to the classical formula (43) for a thick plate. Equations (81) and (82) are compact enough to be used for numerical calculation of the line-shape for any value of R. However, for a general discussion it is convenient to specialize further to the limiting cases $R \gg 1$ and $R \ll 1$.

(c1) *Thick Case, Broad Natural Line*

$$a \gg \lambda \gg 1, \quad R \gg 1.$$

Then, Eq. (75) reduces to

$$L=(1-i)R^{-2}(x+i)^{-1}, \tag{83}$$

and so

$$Z=(\sigma\delta)^{-1}\left[1-i+(1+i)\pi\nu\chi U\left(\frac{f}{1-iU\alpha}-\frac{f'}{1-iU\gamma}\right)\right]. \tag{84}$$

As with Eq. (78), this result is independent of diffusion. But the absorption line is now no longer symmetric, since the real part of Z varies approximately like

$$(1-U\alpha)/(1+U^2\alpha^2), \tag{85}$$

giving the algebraic sum of the real and imaginary parts of the electron susceptibility. That this result (85) holds in the absence of diffusion is well known.[8]

(c2) *Thick Case, Narrow Natural Line*

$$\lambda \gg a \gg 1, \quad R \ll 1.$$

This is the case which will usually occur in practice with bulk metal samples, and it shows the effects of diffusion in the most typical and interesting way. For these reasons we discuss this case in more detail than the others.

To simplify Eq. (82), knowing that $R \ll 1$, it is necessary to consider the center and wings of the absorption line separately. At the center, α is of the order U^{-1}, x is of order 1, and all terms in R^2 in Eq. (82) are negligible. Therefore,

$$L=2R^{-1}(\xi-i\eta)(1+x^2)^{-\frac{1}{2}}. \tag{86}$$

At the wings, α is of the order D^{-1}, and $R^2 x = D\alpha$ is of order 1. In this case Eq. (82) gives

$$L=(D\alpha-1)^{-2}[2(\operatorname{sgn}\alpha-i)|D\alpha|^{-\frac{1}{2}}+(1-i)(D\alpha-3)]. \tag{87}$$

The approximations (86) and (87) overlap and are both valid in the range $U^{-1}<|\alpha|<D^{-1}$. The observed absorption of radio-frequency power per unit area is the quantity P defined by Eqs. (26) and (81). For the center of the line, Eq. (86) gives

$$P=\tfrac{1}{4}\delta\omega|\mathbf{H}_{10}|^2\left[\frac{1}{2\pi}-\nu\chi f(DU)^{\frac{1}{2}}(\operatorname{sgn}\alpha) \times\left(\frac{(1+U^2\alpha^2)^{\frac{1}{2}}-1}{1+U^2\alpha^2}\right)^{\frac{1}{2}}\right], \tag{88}$$

and for the wings Eq. (87) gives

$$P=\tfrac{1}{4}\delta\omega|\mathbf{H}_{10}|^2\{1/2\pi-\tfrac{1}{2}\nu\chi D[f(D\alpha-1)^{-2} \times(2\operatorname{sgn}\alpha|D\alpha|^{-\frac{1}{2}}+D\alpha-3) -f'(D\gamma-1)^{-2}(2(D\gamma)^{-\frac{1}{2}}+D\gamma-3)]\}. \tag{89}$$

The term in γ is small and slowly varying, and has been neglected in Eq. (88) compared with the rapidly varying α term.

Equation (88) is identical with the result (44) obtained from the simple theory of Sec. V. Equation (89)

[8] C. Kittel, Phys. Rev. **73**, 155 (1948), Sec. IV.

shows that the simple theory is incorrect for frequencies far off resonance, even when the sample is thick and the relaxation time long so that the assumption $\lambda \gg a \gg 1$ is valid. However, the complicated and low-intensity wings of the absorption line will be difficult to measure with any accuracy. In the central part of the line, which is all that can be clearly seen in the experiments, the shape will be given by Eq. (88) provided the frequency ω is high enough so that $U \gg D$. As the frequency ω is lowered, D increases and the shape will gradually go over into Eq. (85).

VIII. THEORY INCLUDING SURFACE RELAXATION

The theory so far developed falls short of realism in one respect. The relaxation time U is supposed to take account of the disorientation of electron spins in collisions with lattice vibrations and other electrons. It is likely that such a relaxation time will exist and be roughly independent of position throughout the interior of the metal. However, an electron colliding with the metal surface may be subjected to stronger spin-dependent forces, arising from surface irregularities or paramagnetic surface impurities, than an electron in the interior. An electron colliding with the surface will therefore have a certain probability ϵ of spin disorientation during the collision, in addition to the steady probability U^{-1} per unit time which exists for all electrons. We assume this ϵ to be an average value taken over all collisions with the surface; it is not necessary that the probability be the same for electrons incident at different angles. In this section we shall recalculate the results of Secs. VI and VII, including the effects of surface relaxation as specified by the parameter ϵ.

The effect of the surface relaxation is to change the boundary condition Eq. (9) for the diffusion equation. The probability distribution $F(\mathbf{r},t)$ will now describe the probability of finding, at the point $\mathbf{r}$ at time t, an electron which has not suffered a spin disorientation by surface collision before the time t. Thus $F(\mathbf{r},t)$ describes a diffusion of particles which are lost with probability ϵ every time they hit the metal surface. To find the correct boundary condition, we observe that the mean velocity of the particles at any point is

$$\bar{v} = -\tfrac{1}{3}v\Lambda[\operatorname{grad}F/F], \tag{90}$$

at least when averaged over a time long compared with the collision time τ. Therefore, the number of particles lost per second per unit area of metal surface is

$$\tfrac{1}{3}v\Lambda(\mathbf{n}\cdot\operatorname{grad}F), \tag{91}$$

where $(\operatorname{grad}F)$ is taken at a point just inside the surface, and $\mathbf{n}$ is the inward normal vector. But the number of particles striking the surface per second per unit area is $\frac{1}{4}vF$, and the number lost is by hypothesis $\frac{1}{4}\epsilon vF$. Equating this number to Eq. (91), we find the desired boundary condition

$$\mathbf{n}\cdot\operatorname{grad}F = gF, \quad g = (3\epsilon/4\Lambda). \tag{92}$$

The Green's function G is now given by Eq. (52), where the ψ_n are still solutions of Eq. (48) but with the boundary condition,

$$\mathbf{n}\cdot\operatorname{grad}\psi_n = g\psi_n, \tag{93}$$

instead of Eq. (49). Specializing to the case of a flat plate and solutions even in $(z-\frac{1}{2}\theta)$, we find instead of Eqs. (56) and (57):

$$\psi_n = [\epsilon_n/\theta]^{\frac{1}{2}} \sec u_n \cos[2u_n(z-\tfrac{1}{2}\theta)/\theta], \tag{94}$$

$$\mu_n = 4u_n^2\theta^{-2}, \quad \epsilon_n = 2u_n^2(u_n^2+Q^2+Q)^{-1}, \tag{95}$$

where $u_0, u_1, u_2, \cdots$, are the positive roots of the equation

$$x\tan x = Q, \quad Q = \tfrac{1}{2}g\theta = (3\epsilon\theta/8\Lambda). \tag{96}$$

With these definitions, the results of Sec. VI up to Eq. (59) inclusive are still valid. However, the integration by parts which led to Eq. (60) now gives an additional term which by Eq. (93) reduces to

$$-2g[\epsilon_n/\theta]^{\frac{1}{2}}\mathbf{H}_{10}. \tag{97}$$

The additional term will be carried into the subsequent equations (66), (67) if $(\mathbf{n}\times\mathbf{E}_0)$ is replaced by

$$(\mathbf{n}\times\mathbf{E}_0) + (cg/4\pi\sigma)\mathbf{H}_{10}. \tag{98}$$

Therefore, Eqs. (69) to (71) will be correct with the new definitions of μ_n and ϵ_n, provided Z is replaced by

$$Z + (g/\sigma). \tag{99}$$

The formulas for the line shape are now, instead of Eqs. (71)–(73),

$$Z = (2/\sigma\theta)\{F - Q + \tfrac{1}{2}\pi\nu\chi a^2UF^2[fG(\alpha) - f'G(\gamma)]\}, \tag{100}$$

$$F = [2\sum_0^\infty u_n^2(u_n^2+Q^2+Q)^{-1}(u_n^2-u^2)^{-1}]^{-1}, \tag{101}$$

$$G(\alpha) = -4iu^2\sum_0^\infty u_n^2(u_n^2+Q^2+Q)^{-1} \times (u_n^2-u^2)^{-2}(u_n^2-w^2)^{-1}. \tag{102}$$

The sums can be evaluated in closed form by the following simple physical argument. The value of Z for $\chi = 0$ is obviously independent of Q, since for zero magnetization it does not matter how rapidly the spins are disoriented. Therefore, $(F-Q)$ in Eq. (100) is equal to F given by Eq. (76), and the sum of the series (101) is

$$F = Q - u\tan u. \tag{103}$$

This result can be checked by direct evaluation of Eq. (101) using contour integration. Knowing Eq. (103), we deduce at once from Eq. (95):

$$G(\alpha) = iu(d/du)\{(w^2-u^2)^{-1}[F^{-1}(u) - F^{-1}(w)]\}$$

$$= \frac{i}{(w^2-u^2)^2}\left[\frac{2u^2}{w\tan w - Q} + \frac{w^2-3u^2}{u\tan u - Q} + (w^2-u^2)\frac{u^2\sec^2 u + Q}{(u\tan u - Q)^2}\right]. \tag{104}$$

We next briefly investigate the effects of surface relaxation on the line shapes of Sec. VII.

(a) Thin Case

$$\lambda \leq 4.$$

In this case, Q is small compared to 1, and the smallest root u_0 of Eq. (89) is approximately equal to $Q^{\frac{1}{2}}$. Keeping only the term $n=0$ in the series (101), (102), we find instead of Eq. (78):

$$Z=\frac{2\pi\theta\omega}{c^2}\left[-i+2\pi\nu\chi U\left(\frac{f}{1+\zeta-iU\alpha}-\frac{f'}{1+\zeta-iU\gamma}\right)\right], \quad (105)$$

$$\zeta=(U/T_S), \quad T_S=(2\theta/\epsilon v). \quad (106)$$

The surface relaxation leaves the Lorentz shape of the line unchanged, but increases the width from U^{-1} to $(U^{-1}+T_S^{-1})$. This was to be expected, since T_S is the average relaxation time for an electron travelling back and forth between the two surfaces of the plate and disorienting its spin with probability ϵ at each impact.

For a metal sample of arbitrary shape, and dimensions small compared with the skin depth, Eq. (105) will still hold with

$$T_S=4V/\epsilon vS, \quad (107)$$

where V is the volume and S the surface area of the metal. The physical interpretation of T_S as a relaxation time is the same as before. For example, for a small sphere of radius r,

$$T_S=4r/3\epsilon v. \quad (108)$$

(b) Thick Case

$$\lambda \geq 4, \quad a \geq 4.$$

Putting the approximations $\tan w=\tan u=i$ into Eq. (104), we find that Eq. (81) holds with

$$L=[R^2(x+i)-1]^{-2}\left[\frac{(2+(1+i)R\psi)^2}{R(\xi+i\eta+i\psi)}-2R\psi+(1-i)[R^2(x+i)-3]\right], \quad (109)$$

where we have written

$$\psi=(2Q/a)=\epsilon(3U/8\tau)^{\frac{1}{2}}. \quad (110)$$

The ratio (τ/U) is the probability of spin disorientation for an electron in a single scattering collision inside the metal, and in Li and Na may be of the order of 10^{-5}, while ϵ is the corresponding probability for a collision at the surface and may be larger by a considerable factor. We do not know *a priori* whether ψ will be large or small compared to 1.

(b1) *Thick case, broad natural line*

$$a \gg \lambda \gg 1, \quad R \gg 1.$$

In this case, the ψ-dependent terms in Eq. (109) are negligible and we still have the results (84) and (85).

(b2) *Thick case, narrow natural line*

$$\lambda \gg a \gg 1, \quad R \ll 1.$$

In this case, Eq. (109) gives for the central part of the line

$$L=4R^{-1}(\xi-i\eta-i\psi)[\xi^2+(\eta+\psi)^2]^{-1}. \quad (111)$$

Thus the central part of the observed absorption line has the shape given by

$$P=\tfrac{1}{4}\delta\omega|\mathbf{H}_{10}|^2\{(1/2\pi)-\nu\chi f(DU)^{\frac{1}{2}}(\mathrm{sgn}\alpha)\times[(1+U^2\alpha^2)^{\frac{1}{2}}-1]^{\frac{1}{2}}[(1+U^2\alpha^2)^{\frac{1}{2}}+\psi\{(1+U^2\alpha^2)^{\frac{1}{2}}+1)^{\frac{1}{2}}+\tfrac{1}{2}\psi^2]^{-1}\}. \quad (112)$$

If $\psi \ll 1$, this reduces to Eq. (88). For $\psi \gg 1$, the line has a different limiting shape:

$$P=\tfrac{1}{4}\delta\omega|\mathbf{H}_{10}|^2\{(1/2\pi)-2\nu\chi f\psi^{-2}(DU)^{\frac{1}{2}}(\mathrm{sgn}\alpha)\times[(1+U^2\alpha^2)^{\frac{1}{2}}-1]^{\frac{1}{2}}\}, \quad (113)$$

but is still approximately antisymmetrical about its center.

For making precise comparisons of theory with experiment in the thick case, the exact Eq. (109) must be used. However, it seems likely, from the rather sharp lines observed in experiments with small particles of metal,[9] that ψ is small compared to 1. If this is so, then the only effect of surface relaxation is a possible broadening of the lines in the thin case, and the results of Sec. VII will apply to the thick case without serious error.

IX. CONCLUSIONS

The results of this analysis may be summarized as follows:

(1) For thin samples ($\theta<4\delta$), the absorption line has the usual symmetrical shape $[1+\alpha^2\bar{U}^2]^{-1}$, with a width $\bar{U}^{-1}=U^{-1}+T_S^{-1}$, where U and T_S are the volume and surface relaxation times.

(2) For thick samples the line has a central structure of width equal to the natural width U^{-1}, with wings extending over a band D^{-1} if $D<U$.

(3) The central structure is always markedly more intense than the wings, so that under normal experimental circumstances the apparent width of the line will be of the order of U^{-1}.

(4) The characteristic effect of electron diffusion is not to broaden the line but to make a radical change in its shape.

(5) For thick samples with a broad natural width

[9] G. Feher and A. F. Kip, reference 1, Appendix A.

($\theta \gg \delta$, $D \gg U$), the line has the shape

$$[(1-\alpha U)/(1+\alpha^2 U^2)].$$

(6) For thick samples with a narrow natural width ($\theta \gg \delta$, $U \gg D$), and not too strong surface relaxation ($\psi \ll 1$), the central part of the line has the antisymmetrical shape

$$\mathrm{sgn}\alpha\left(\frac{(1+U^2\alpha^2)^{\frac{1}{2}}-1}{1+U^2\alpha^2}\right)^{\frac{1}{2}}. \qquad (114)$$

(7) For thick samples, the intensity of the line at the center is reduced by the diffusion effect by a factor of the order of $(D/U)^{\frac{1}{2}}$. When $D \ll U$, the integrated intensity of the line comes mainly from the diffuse wings and not from the center.

(8) Results (1)–(7) are independent of the shape of the metal, although the detailed calculations of Secs. VI–VIII apply only to flat plates.

Experimentally, it is found convenient to measure not the actual energy absorption for a given **H** but the change in absorption resulting from a small modulation of **H**. Thus the measured quantity is the derivative of P with respect to ν. This has the effect of further accentuating the center of the line as compared with the wings. In particular, the shape (114) appears as a very strong symmetrical peak with shallow minima (depth about 6 percent of the central peak) on either side.

The author wishes to express his thanks to Professor Kittel for suggesting this problem to him, to the whole of the Berkeley solid-state physics group for their active cooperation, and to the Department of Physics of the University of California for its hospitality during the time this work was done.

PHYSICAL REVIEW VOLUME 98, NUMBER 2 APRIL 15, 1955

Oscillatory Thermomagnetic Properties of a Bismuth Single Crystal at Liquid Helium Temperatures

M. C. Steele and J. Babiskin
United States Naval Research Laboratory, Washington, D. C.
(Received December 16, 1954)

Measurements of the thermoelectric power, thermal conductivity, and electrical resistance have been made on a bismuth single crystal in magnetic fields up to 13 kilogauss at liquid helium temperatures. The results for two newly discovered effects, namely the oscillatory magnetic field dependence of both the thermoelectric power and the thermal conductivity, are presented in detail. The crystal was mounted transverse to the field with the trigonal axis parallel to the crystal length. The heat current was also parallel to the crystal length. Most of the data were taken with the magnetic field parallel to one of the crystalline binary axes where only one oscillating component was found. For this orientation the oscillations in the two thermomagnetic properties were periodic in H^{-1} with a period of 7.1×10^{-5} gauss^{-1}. This result is compared to the calculated period obtained from susceptibility oscillations (de Haas-van Alphen effect) of bismuth and found to agree very well. Temperature and field dependence of the amplitude of the oscillatory effects are discussed.

INTRODUCTION

THE thermal and electrical properties of bismuth have a long history of peculiar behavior especially when placed in a magnetic field. These peculiarities are emphasized when low temperatures are combined with strong magnetic fields. Thus, the pioneer work of de Haas and van Alphen[1] on the oscillatory susceptibility of Bi opened the door to susceptibility, magnetoresistance, and Hall effect experiments on many other metal single crystals. The latter two properties fall in the catagory of galvanomagnetic effects. Such measurements have been made on Bi and have revealed that both the magnetoresistance[2] and Hall coefficient[3] exhibit the same type of magneto-oscillatory behavior as the susceptibility. However, there has been only a little work done[4,5] on the thermomagnetic properties of Bi single crystals at low temperatures.

The present investigation was undertaken to determine if the thermoelectric power and thermal conductivity of Bi exhibit the same type of magneto-oscillatory behavior as do the susceptibility, magnetoresistance, and Hall coefficient. On the basis of thermodynamic arguments relating the galvanomagnetic and thermomagnetic coefficients,[6] it was believed that such properties as the thermoelectric power and thermal conductivity would show the oscillatory dependence on the magnetic field strength. Bismuth was chosen as the metal to study since its magnetic behavior is much more pronounced than most other metals at

[1] W. J. de Haas and P. M. van Alphen, Leiden Comm. **212A** (1930).
[2] P. B. Alers and R. T. Webber, Phys. Rev. **91**, 1060 (1953).
[3] Reynolds, Leinhardt, and Hemstreet, Phys. Rev. **93**, 247 (1954). See also Laird C. Brodie, Phys. Rev. **93**, 935 (1954).
[4] S. Shalyt, J. Phys. (U.S.S.R.) **8**, 315 (1944).
[5] de Haas, Gerritsen, and Capel, Physica **3**, 1143 (1936).
[6] H. B. Callen, Phys. Rev. **85**, 16 (1952).

low temperatures. Preliminary reports[7,8] on this work have already been given.

EXPERIMENTAL DETAILS

(A) Specimen

The single crystal used throughout all the experiments reported here was grown from Cerro de Pasco Copper Company bismuth (99.99 percent pure). Since a particular orientation was desired (trigonal axis parallel to crystal length), it was grown by a seeding technique in a furnace described previously by Schubnikow.[9] The crystal, as used, was a cylinder 57 mm long and 1.7 mm in diameter. The trigonal axis had been located by cleaving the crystal under liquid nitrogen. From goniometer measurements it was found that the trigonal axis was parallel to the crystal length to within 1.5°.

(B) Magnet

Horizontal magnetic fields, H, up to 13 kilogauss were supplied by an A. D. Little electromagnet capable of rotation about a vertical axis. The 11-inch diameter pole pieces were separated by a gap of 4 inches. With this magnet arrangement and the relatively small size of the crystal, homogeneous magnetic fields were readily attained. Calibration of the magnet, to within 0.02 percent, was accomplished through use of a nuclear fluxmeter.

(C) Measurement Technique

A noninductive manganin heating coil on a coil form was soldered to the upper and lower ends of the bismuth crystal. The crystal with the heaters was mounted vertically from the top plate of a vacuum can, so that the upper heater form was in thermal contact with the liquid helium bath. The vacuum can was then sealed with a gold ring seal. The lower heater established the temperature gradient. The upper heater established the mean temperature of the crystal at a temperature slightly above the bath temperature. With this arrangement, the effects of a drifting bath temperature and of magnetic variations in the thermal conductivity of the crystal could be compensated by varying the current in the upper heater. All the heater and thermometer leads were number 40 formex insulated manganin wires and were wrapped on and glyptalled to the top plate in order to bring them into thermal contact with the liquid helium bath. The leads were strung through a sharply bent pumping tube which served as a radiation shield and were brought out of the vacuum system through a multiple-terminal Stupakoff glass to metal seal at room temperature. The vacuum system incorporated a provision for the addition of helium exchange gas into the vacuum can for isothermal measurements such as magnetoresistance.

Thermoelectric junctions to copper wire (No. 44) were made at points 47 mm apart on the crystal. These copper leads were brought out to the liquid helium bath through Stupakoffs in the top plate of the vacuum can. Thus, for all the experiments to be reported here, H was perpendicular to the trigonal axis of the crystal, while the heat was parallel to the trigonal axis. This is the desired orientation because susceptibility oscillations are then more pronounced than if H were parallel to the trigonal axis. The binary axes of the crystal were located from magnetoresistance rotation measurements at 4.2°K. For the actual field used the resistance was maximum when H was perpendicular to a binary axis and minimum when H was parallel to a binary axis. Carbon resistance thermometers were placed in copper jackets which in turn were soldered to the crystal diametrically opposite the thermoelectric junctions. A low-melting (55°C) indium solder was used to make all connections to the crystal. The carbon thermometers were calibrated for both temperature and magnetic field, so that the small resistance corrections due to the magnetoresistance of the carbon thermometer could be applied.

A thermofree six-dial potentiometer was used to measure the thermal emf. Less precise potentiometers were used to measure the voltage drop across the carbon resistors which were connected in series and carried a constant measuring current ($\sim 10^{-5}$ amp). The input measuring power was $\sim 10^{-7}$ watt. Heater currents were of the order of milliamperes and could be read so that the error was less than 0.02 percent. The power supplied by the heaters was of the order of milliwatts.

A typical run proceeded in the following manner. The temperature gradient and upper junction temperature were both fixed by adjusting the heaters in zero magnetic field. After bringing H up to maximum value, the heaters were readjusted to give the same temperature conditions as in zero field. For every decrease of the magnetic field corresponding to an experimental point, the potentiometers across the carbon thermometers were first adjusted to correct for the magnetoresistance, then the current in the upper and lower heaters was adjusted to satisfy the foregoing conditions. In this way, both the temperature gradient and the upper temperature were maintained constant throughout a field sweep. The thermoelectric voltage and heater currents were then recorded as a function of field as H was decreased. This procedure gave simultaneous measurements of thermoelectric power and thermal conductivity as a function of magnetic field, with temperature as a parameter.

The vacuum in the can during an experimental run was $\sim 10^{-2}\mu$ of mercury as read on a Phillips gauge at room temperature. Actually it is presumed that the vacuum in that part of the system at liquid helium temperatures was considerably better than the above value.

[7] M. C. Steele and J. Babiskin, Phys. Rev. **94**, 1394 (1954).
[8] J. Babiskin and M. C. Steele, Phys. Rev. **96**, 822 (1954).
[9] L. W. Schubnikow, Leiden Comm. **207B** (1930).

EXPERIMENTAL RESULTS

(A) Thermoelectric Power

In Fig. 1. are plotted curves for thermoelectric power as a function of H^{-1} for three different temperatures. Relative to the sign convention all the thermoelectric powers reported in this paper are *negative*. For this set of data, H was parallel to one of the binary axes. The curves show clearly the oscillatory nature of the thermoelectric power even at the highest temperature (4.33°K). These measurements were made on the same day so as to observe the gross behavior of the thermoelectric power as a function of magnetic field and temperature. Actually these data were taken on an $X-Y$ recorder (with a preamplifier) in order to be able to get sufficient data in a given helium run. For these data the heater currents were constant, so that the variation of ΔT could have been ~ 2 percent over the entire field sweep due to the magnetic dependence of the thermal conductivity. The thermal voltages ranged up to 80 microvolts and were recorded as the Y coordinate, while the magnet current was the X coordinate. The points shown in Fig. 1 were taken from the $X-Y$ trace so as to have equal increments in H^{-1}. These data (Fig. 1) are not as precise as those obtained with the six-dial potentiometer.

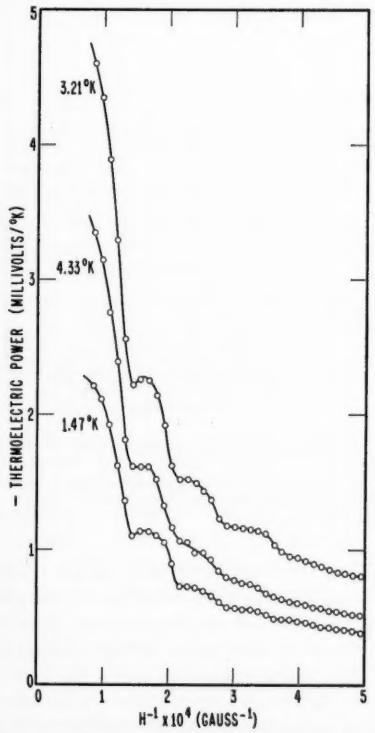


FIG. 1. Thermoelectric power as a function of H^{-1} for a bismuth single crystal at three temperatures. These data were taken at the mean temperatures of 3.21°K, 4.33°K, and 1.47°K, with temperature differences of 0.017°K, 0.028°K, and 0.033°K respectively between thermojunctions.

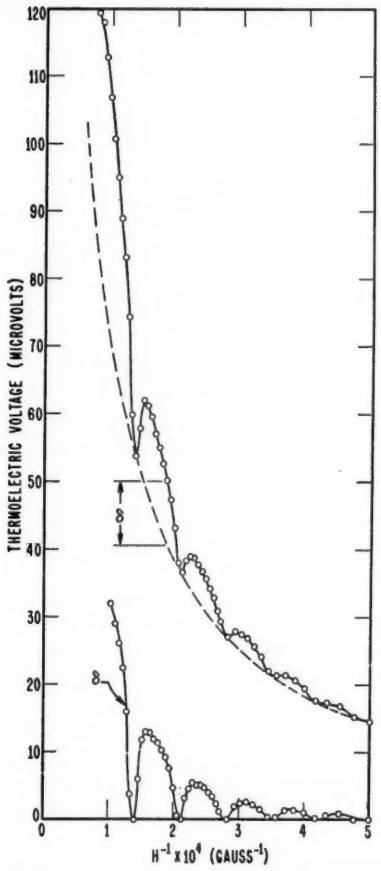


FIG. 2. Thermoelectric voltage as a function of H^{-1} when H is parallel to a binary axis of the bismuth crystal. The mean temperature was 1.604°K and the temperature difference was 0.064°K. δ is the difference between the experimental points and the dashed envelope curve.

Figure 2 shows the results (with the same orientation as for Fig. 1) at a mean temperature of 1.604°K and a temperature difference of 0.064°K between thermojunctions. To obtain the thermoelectric power in microvolts per degree, one needs only to divide the voltages in Fig. 2 by the constant temperature gradient, 0.064°K. These data were obtained with the six-dial potentiometer. The error in these voltage readings was of the order of 0.5 percent. In order to emphasize the oscillatory component we have also plotted in Fig. 2 the difference, δ, between the experimental points and the dashed envelope curve which passes through the minima of the data. We shall show later that the δ curve is strictly periodic in H^{-1}. At this point it is noted that the amplitude of the oscillatory component becomes monotonically larger as we go to higher values of H. At $H^{-1} \simeq 1.6 \times 10^{-4}$ gauss^{-1}, corresponding

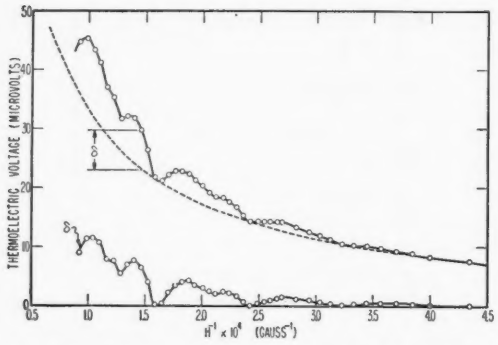


FIG. 3. Thermoelectric voltage as a function of H^{-1} when H is perpendicular to a binary axis of the bismuth crystal. The mean temperature was 1.678°K while the temperature difference was 0.022°K. δ is the difference between the experimental points and the dashed envelope curve. Note the structure in the δ curve here as compared to that in Fig. 2. This is due to at least two frequencies of oscillation being present.

to a maximum value for δ, δ (which is then twice the oscillatory amplitude) is 27 percent of the envelope value.

Figures 3 and 4 show the thermoelectric results as a function of H^{-1} for two different orientations. These results are presented to exemplify the existence of beats resulting from more than one oscillating term. In Fig. 3, H was perpendicular to a binary axis, while in Fig. 4 H made an angle of 45° with a binary axis. Since these results are more complicated than those shown in Figs. 1 and 2 we shall confine our attention in this paper to the case with H parallel to a binary axis. A detailed study of the other orientations would require another investigation.

(B) Thermal Conductivity

Oscillations in the thermal conductivity as a function of magnetic field were observed clearly only in the lowest temperature (mean temperature 1.604°K) run.

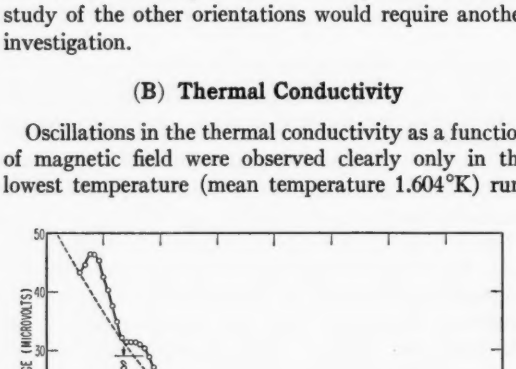


FIG. 4. Thermoelectric voltage as a function of H^{-1} when H is at 45° to a binary axis of the bismuth crystal. The mean temperature was 4.370°K and the temperature difference was 0.014°K. δ is the difference between the experimental points and the dashed envelope curve. Again there is evidence of more than one frequency of oscillation.

At higher temperatures (up to ~5°K for the mean temperature) the results only suggested the possibility of oscillations. Therefore we present only the data for the lowest temperature. In Fig. 5 the thermal conductivity is plotted as a function of H^{-1}. Here too H was parallel to a binary axis. The dashed curve is the envelope passing through minima while the solid curve traces the experimental points. Differences δ between the envelope and the experimental points are also plotted in Fig. 5. The oscillatory character of the thermal conductivity is clearly evident from these results. These data were obtained simultaneously with the thermoelectric data shown in Fig. 2, so that we

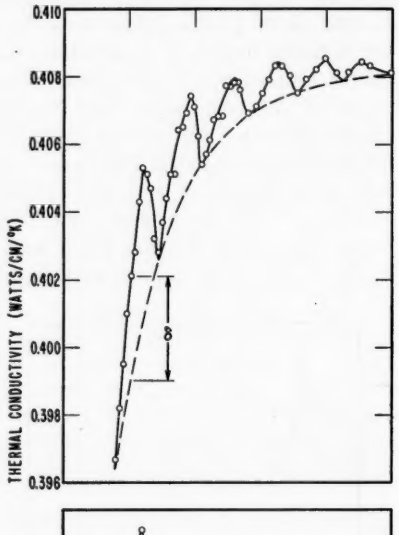


FIG. 5. Thermal conductivity as a function of H^{-1}. H was parallel to a binary axis of the bismuth crystal. The mean temperature was 1.604°K and the temperature difference was 0.064°K. δ is the difference between the experimental points and the dashed envelope curve.

shall be able to compare details of the oscillations readily. Thus, the mean temperature for the results in Fig. 5 was 1.604°K and the temperature difference was 0.064°K. The zero field value of the thermal conductivity at this temperature was 0.415 watt units. The amplitude of the oscillations in δ becomes monotonically larger as H is increased. At $H^{-1} \simeq 1.2 \times 10^{-4}$ gauss^{-1}, corresponding to a maximum value for δ, δ is only 1 percent of the envelope value. The relative error for this thermal conductivity data was less than 0.1 percent. The mean value of the thermal conductivity

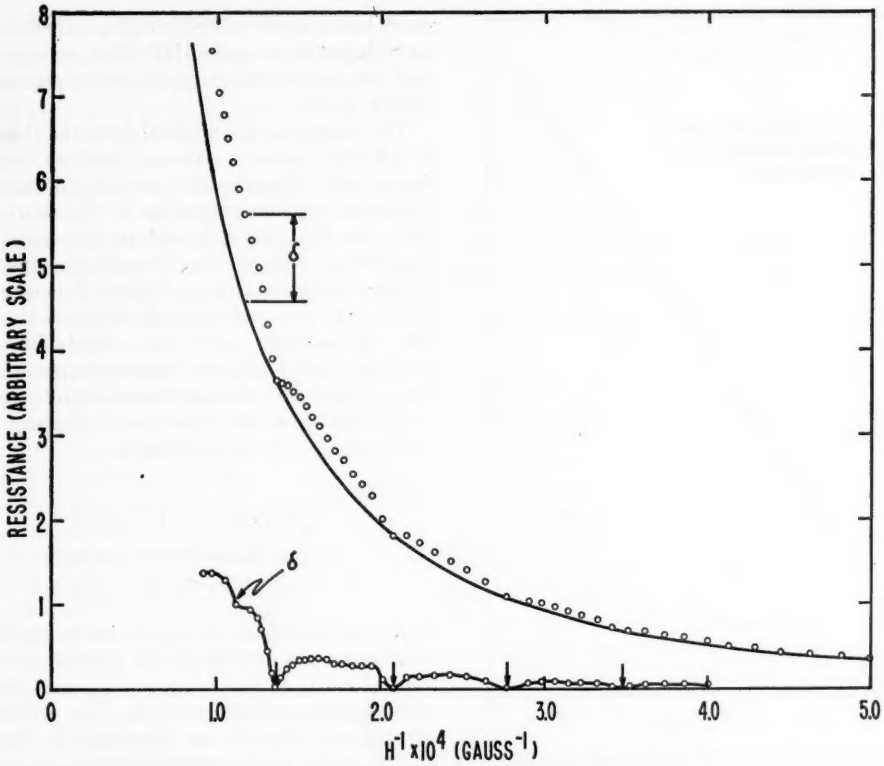


FIG. 6. Electrical resistance of a bismuth single crystal as a function of H^{-1} at $T=1.232°K$. H was parallel to a binary axis and perpendicular to the trigonal axis. The electric current was parallel to the trigonal axis. δ is the difference between the experimental points and the solid envelope curve.

decreased by approximately 4 percent in going from zero field to the maximum H. So it is clear that for Bi the thermoelectric power oscillations with magnetic field are far easier to observe experimentally than are the thermal conductivity oscillations. In fact, the smallness of the effect in thermal conductivity is probably one of the reasons it was not observed in previous measurements[4] of thermal conductivity as a function of magnetic field.

(C) Magnetoresistance

Some magnetoresistance measurements were made on this crystal so as to be able to correlate the various magneto-oscillatory effects on the same crystal. No systematic study of magnetoresistance was needed because of the comprehensive work of Alers and Webber.[2]

Our magnetoresistance results at 1.232°K are shown in Fig. 6. H was parallel to a binary axis and perpendicular to the trigonal axis, while the electric current was parallel to the trigonal axis. For these data the potential drop across the specimen was recorded directly as a function of magnet current on the $X-Y$ recorder. Points shown in Fig. 6 were taken from the continuous $X-Y$ trace. The solid curve is an envelope passing through minima. δ, the difference between the envelope and experimental points, is plotted against H^{-1} at the bottom of Fig. 6. The arrows on the δ curve denote values for H^{-1} corresponding to resistance minima. The amplitude of these oscillations in δ also exhibits a monotonic increase with increasing H. At $H^{-1}\simeq 1.6\times 10^{-4}$, gauss^{-1}, corresponding to a maximum in δ, δ is 13 percent of the envelope value. The error in this data was less than 0.5 percent.

DISCUSSION

(A) Period of Oscillations

The correlation of these oscillatory thermomagnetic effects to the oscillatory susceptibility and galvanomagnetic effects is most readily shown by comparing the period of oscillation in H^{-1}. In Landau's[10] theory of the de Haas–van Alphen effect this period is given by β/E_0, where β is a double effective Bohr magneton and E_0 is the chemical potential (Fermi energy) of the pertinent de Haas–van Alphen electrons (or holes) measured from the bottom of the significant zone for electrons (or from the top of the zone for holes). So we

[10] L. D. Landau, see appendix to D. Schoenberg, Proc. Roy. Soc. (London) A170, 341 (1939).

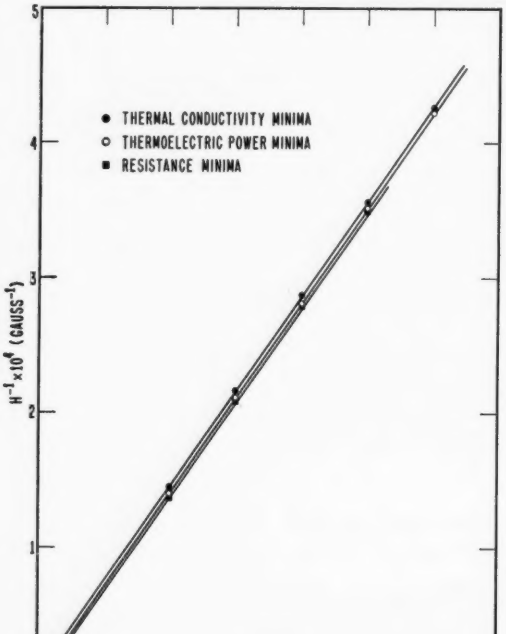


FIG. 7. Values of H^{-1} at which thermal conductivity, electrical resistance, and thermoelectric power minima occur in a bismuth single crystal, plotted against integers. The data for this plot were taken from Figs. 2, 5, and 6.

must first show that the oscillations in the thermoelectric power and thermal conductivity are periodic in H^{-1}. This can be seen clearly from Fig. 7 where the values of H^{-1} corresponding to minima in the δ curves of Figs. 2, 5, and 6 are plotted against integers. A straight line in such a plot is proof of the periodicity in H^{-1}. The three straight lines in Fig. 7 show that for this Bi crystal the thermoelectric power, thermal conductivity, and magnetoresistance are all periodic in H^{-1}. Furthermore, the fact that the lines are all parallel (to within experimental error) indicates that they have the same periodic of oscillation, β/E_0. Our assignment of the value β/E_0 to the period of oscillation for the thermomagnetic effects presupposes that when a proper theory for these effects is formulated such a result will be part of it. In the case of magnetoresistance there is some theoretical basis[11,12] for assigning the value β/E_0 to the period. A very interesting conjecture on the fact that all the transport properties of Bi should exhibit magneto-oscillatory effects with the same period as the susceptibility was made by Robinson[13] several years before there existed complete experimental proof on such properties as the Hall effect, magnetoresistance and the two thermomagnetic effects discussed in the present paper.

The value of β/E_0 obtained from the lines in Fig. 7 is 7.1×10^{-5} gauss^{-1}. We now seek to compare this value with Shoenberg's[14] results on susceptibility. Unfortunately, our orientation for the above value was with the field perpendicular to a trigonal axis and parallel to a binary axis. For such an orientation the couple, measured on a torsion balance, vanishes. Hence, we must pass to a calculation of the period for the susceptibility with this orientation from the constant energy ellipsoid scheme proposed by Shoenberg. Using the revised parameters given by Shoenberg,[15] we find that there are three (two of which are the same) values of β/E_0 for our orientation.

They are

$$\beta_1/E_0=2.5\times10^{-6}\ \text{gauss}^{-1},$$
$$\beta_2/E_0=7.1\times10^{-5}\ \text{gauss}^{-1},$$
$$\beta_3/E_0=7.1\times10^{-5}\ \text{gauss}^{-1}.$$

The values for β_2/E_0 and β_3/E_0 are in excellent agreement with the results of the present work. This is further evidence for the generality and similarity of the magneto-oscillatory effects. The absence of any short period term in our observations, corresponding to the β_1/E_0 of the susceptibility calculation, is not too disturbing or surprising since, in general, the amplitude of short period oscillations is much less than that of longer period ones.

For H perpendicular to both a binary axis and the trigonal axis, the calculated periods for susceptibility oscillations are

$$\beta_1/E_0=8.2\times10^{-5}\ \text{gauss}^{-1},$$
$$\beta_2/E_0=4.1\times10^{-5}\ \text{gauss}^{-1},$$
$$\beta_3/E_0=4.1\times10^{-5}\ \text{gauss}^{-1}.$$

Although our results (Fig. 3) indicate that there are at least two oscillating terms present, we cannot readily make a quantitative comparison between calculated and observed periods.

The same situation pertains to the orientation with H at a 45° angle to a binary axis (and H still perpendicular to the trigonal axis). In that case the calculated periods are

$$\beta_1/E_0=5.8\times10^{-5}\ \text{gauss}^{-1},$$
$$\beta_2/E_0=7.9\times10^{-5}\ \text{gauss}^{-1},$$
$$\beta_3/E_0=2.1\times10^{-5}\ \text{gauss}^{-1}.$$

The results in Fig. 4 confirm the existence of more than one term but, again, a quantitative check is not fully justified.

[11] A. Achieser, Compt. rend. **23**, 874 (1939).
[12] B. Davydov and I. Pomeranchuk, J. Phys. (U.S.S.R.) **2**, 147 (1940).
[13] J. E. Robinson, dissertation, Yale University, 1950 (unpublished).
[14] D. Schoenberg, Proc. Roy. Soc. (London) **A178**, 341 (1939).
[15] D. Schoenberg, Trans. Roy. Soc. (London) **A245**, 1 (1952).

(B) Relative Phase of the Oscillations

The relative phase of the various oscillations is of interest from a theoretical point of view. Such information can aid in qualitative interpretation of how the magnetic field is affecting each of the properties.

From Fig. 7, we see that minima in thermal conductivity, thermoelectric power, and electrical resistance occur at very nearly the same field strength. Actually there are small phase differences between the three lines in Fig. 7. In units of π radians, these differences are 0.16 between thermal conductivity and thermoelectric power, and 0.08 between thermoelectric power and electrical resistance. It is clear that the relative phase we are reporting here is dependent on the way the envelope curves were drawn. Thus the small differences found in Fig. 7 may be within the limits imposed by a variety of envelope curves that could be put through the experimental points.

If one assumed that the Lorenz number (product of the thermal conductivity and the electrical resistivity divided by the temperature) remained constant for Bi at these low temperatures and relatively strong magnetic fields, then one would have expected a maximum in thermal conductivity to correspond to a minimum in electrical resistance. Since we do not observe this, we conclude that the crystal lattice contribution to the thermal conduction is not independent of the magnetic field. A mechanism that involves the lattice in thermal conductivity of metals is the scattering of lattice vibrations by electrons. Such scattering has been shown[16] to produce a term in the thermal conductivity $\propto 1/N_a^2$, where N_a is the number of free electrons per atom. On the other hand, from elementary concepts we have that the electrical resistance $\propto 1/N_a$. If one now assumes that N_a varies with magnetic field strength, it is found that a minimum in electrical resistance would indeed occur at the same field strength as a minimum in the thermal conductivity. Although the above argument is only qualitative (particularly since the $1/N_a^2$ dependence for the electron-lattice vibration scattering may no longer be valid at low temperatures) it is presented in any event since there is, at present, no adequate rigorous theory to explain the observations.

Further, the phase relation between the observed thermoelectric power and magnetoresistance oscillations cannot be explained in light of existing theory, since the only model for which thermomagnetic and galvanomagnetic calculations have been made is the two-band isotropic metal.[17] It is noted though that the occurrence of minima in the oscillatory component of the thermoelectric power and magnetoresistance at nearly the same field is in keeping with the fact that the envelope curves of both these quantities also show similarities.

[16] J. L. Olsen and H. M. Rosenberg, Advances in Physics **2**, 28 (1953).

[17] A. H. Wilson, *Theory of Metals* (Cambridge University Press, New York, 1953), p. 208.

TABLE I. Temperature dependence of thermoelectric power for $H=5500$ gauss.

T_{mean}°K	ϵ_m, mv/°K	$\Delta\epsilon$, mv/°K	$\Delta\epsilon/\epsilon_m$
1.47	0.99	0.12	0.121
3.21	1.96	0.18	0.092
4.33	1.41	0.11	0.078

Although both increase with increasing H, we shall show later that the envelope curves do not follow the same field dependence quantitatively.

(C) Temperature Dependence of the Thermoelectric Power

Since the thermoelectric oscillations were readily observed at the three temperatures shown in Fig. 1, we can discuss the temperature dependence of both the oscillations and the mean value of the thermoelectric power. It is clear that the 3.21°K data have the largest thermoelectric power, ϵ, at all field strengths. But this does not give the complete picture insofar as the oscillations are concerned. In particular, we shall be interested in knowing the relative magnitude of the oscillations compared to the mean value at any field. To illustrate such a temperature dependence, there are shown in Table I data taken from Fig. 1 at $H=5500$ gauss. The mean value of the thermoelectric power is designated as ϵ_m, while the amplitude of the oscillatory component is $\Delta\epsilon$. The field value chosen corresponds to a maximum in the oscillatory component at each temperature.

Table I exhibits several points of interest. The ratio, $\Delta\epsilon/\epsilon_m$, becomes larger as the temperature is decreased. This means that at lower temperatures the oscillatory character of the thermoelectric power is more readily extracted from the total thermoelectric power. But there is probably no great experimental advantage to going to much lower temperatures that 1.47°K because the value of ϵ_m has already decreased by a factor of two in going from 3.21°K to 1.47°K. Thus, one would need to double the temperature gradient at 1.47°K in order to observe the same thermal voltage as measured at 3.21°K. It seems reasonable to expect that ϵ_m will continue to drop with decreasing temperatures like T^n, where n is of order 1. On the other hand, it may be of interest to make measurements at temperatures higher than 4.33°K since the oscillations still amounted to 7.8 percent of the total at this temperature and a field of only 5500 gauss. Higher field and temperature measurements might reveal additional clues to aid in the explanation of the phenomena. In fact, we have also made some measurements (for the same crystal) on the field dependence of ϵ at liquid nitrogen temperatures. These results, although preliminary in nature, indicate a rather marked and interesting field dependence. Thus, at 50°K, with H parallel to a binary axis (and still perpendicular to the trigonal axis) ϵ exhibited a maximum at about 6 kilogauss. Measurements were

TABLE II. Field dependence of the maxima in the oscillations of thermoelectric power ($T_{mean}=1.604°K$).

$H^{-1}\times10^4$ (gauss^{-1})	$\delta(\epsilon)$, μv	$\mathcal{E}$, μv	$\delta(\epsilon)/\mathcal{E}$
1.6	13.0	47.8	0.272
2.3	5.4	33.3	0.162
3.1	2.6	24.4	0.106
3.8	1.6	19.4	0.083
4.5	1.0	16.1	0.062

continued to 13 kilogauss and showed that ϵ was still decreasing. When H was at 45° to a binary axis (at 50°K) ϵ showed a maximum at 1.9 kilogauss and a minimum at 6.7 kilogauss. For this orientation ϵ was still increasing at 13 kilogauss. Such behavior for ϵ suggests the possibility of a magneto-oscillatory thermomagnetic effect in Bi at comparatively high temperatures. A more complete study of this behavior is contemplated in order to clarify the possibility suggested above.

(D) Field Dependence of Envelope and Amplitude of Oscillations

(a) *Thermoelectric Power*

One of the questions that must be answered by any complete theory of thermomagnetic phenomena is the manner in which the effects depend on magnetic field strength. For bismuth this explanation would need to encompass the field dependence of both the oscillation amplitude and the envelope curve. Figure 2 gives such information on the thermoelectric power at a mean temperature of 1.604°K. Over the field range investigated, the dashed line (envelope curve) in Fig. 2 is represented by

$$\epsilon \propto H^{1.0}. \tag{1}$$

Similar analysis of the data given in Fig. 1 show that this field dependence for ϵ is very nearly temperature independent (over the range 1.47°K–4.33°K).

Some idea of the field dependence of the oscillations in ϵ can be obtained by noting how maxima of the δ curve (Fig. 2) change with H. The pertinent information is given in Table II. $\mathcal{E}$ designates the envelope value.

Analysis of the data in Table II shows that the relation

$$\delta(\epsilon) \propto H^{2.5} \tag{2}$$

describes the field dependence of the oscillation amplitude very well. Such a field dependence differs markedly from that exhibited by the susceptibility[14] oscillations where the amplitude either shows a leveling off or actually attains a maximum value as H is increased. Of course one needs to extend the thermoelectric measurements to stronger fields in order to see if Eq. (2) continues to represent the data correctly. The last column in Table II, $\delta(\epsilon)/\mathcal{E}$, shows that the oscillations become far easier to observe (relative to the total thermoelectric power) as H is increased.

To complete this discussion we should like to give an estimate of how many fold ϵ is increased from its zero field value to its value at $H=10$ kilogauss. To make this estimate we note from Fig. 1 that at $H=10$ kilogauss and $T_m=4.33°K$, $\epsilon_m\simeq3\times10^{-3}$ volt/°K. In an earlier work on polycrystalline Bi wire, we reported[18] a zero-field value of $\epsilon\simeq6\times10^{-6}$ volt/°K at 5°K. Since it was also noted there that the zero field value of ϵ for a Bi single crystal was only somewhat larger than for the polycrystalline wire, we feel justified in using this estimate to obtain the desired ratio. Thus ϵ is increased by about a factor of 500 in going from zero field to 10 kilogauss.

(b) *Thermal Conductivity*

From Fig. 5 we find that the envelope curve (dashed line) of the thermal conductivity, κ, is practically field independent over the range of fields investigated. In fact, we have already noted that κ_{mean} only decreased by about 4 percent in going from zero field to 13 kilogauss. On the other hand, the amplitude of the oscillations in κ are strongly field dependent. Data on how the maxima of the $\delta(\kappa)$ curve change with H is given in Table III.

TABLE III. Field dependence of the maxima in the oscillations of thermal conductivity ($T_{mean}=1.604°K$).

$H^{-1}\times10^4$ (gauss^{-1})	$\delta(\kappa)$, watt units
1.2	0.0046
1.9	0.0026
2.6	0.0014
3.3	0.0010
4.0	0.0007

From Table III it is found that

$$\delta(\kappa)\simeq H^{2.0} \tag{3}$$

describes the field dependence of the oscillation amplitude over the range investigated. Here again, the field dependence is different from that shown by the susceptibility.

(c) *Electrical Resistance*

For the sake of completeness on this crystal we also give the field dependence of the electrical resistance ρ. Data (from Fig. 6) on the field dependence of the envelope curve and the maxima in $\delta(\rho)$ are given in Table IV.

Analysis of these data shows that the envelope curve can be described by the relation

$$\rho \propto H^{2.0}, \tag{4}$$

but the oscillation amplitude field dependence is more complicated than a simple power of H. Equation (4) is in reasonably good agreement with the results of

[18] J. Babiskin and M. C. Steele, Phys. Rev. 95, 602 (1954).

Alers and Webber[2] who found $\rho \propto H^{1.66}$ for a different orientation (H perpendicular to both a binary axis and the trigonal axis).

(E) Separation of Thermomagnetic Effects into Even and Odd Field Components

In Fig. 8 are shown data of the thermoelectric power (arbitrary units) as a function of magnet angle for $H=12\,600$ gauss, $T_m=5.298°K$, and $\Delta T=0.128°K$. For a magnet angle of 0°, H was perpendicular to a binary axis. The field was perpendicular to the trigonal axis throughout the entire rotation.

These data are given to show the lack of 180° symmetry in the rotation of H. The same lack of symmetry was noted in the thermal conductivity measurements at 1.604°K. But our measurements on the electrical resistance in a magnetic field (a galvanomagnetic property) showed that $\rho(H)=\rho(-H)$ very precisely. Earlier measurements of thermomagnetic

TABLE IV. Field dependence of the envelope and oscillation amplitude of electrical resistance. ($T=1.232°K$.)

$H^{-1}\times10^4$ (gauss^{-1})	ρ (arbitrary units)	$\delta(\rho)$ (arbitrary units)	$\delta(\rho)/\rho$
1.6	2.82	0.36	0.128
2.4	1.40	0.17	0.121
3.1	0.86	0.09	0.105
3.8	0.57	0.05	0.088

effects in Bi crystals at liquid nitrogen temperatures by Grüneisen and Gielessen[19,20] had also shown such an "umkehreffekt." Later work by Grüneisen *et al.*[21] on Bi showed that the effect was still there but that the explanation was still lacking. Similar results have also been obtained by Rausch[22] in his thermomagnetic measurements on antimony crystals at liquid nitrogen temperatures.

Further data on the effect of field reversal were taken at lower temperatures, and revealed that the lack of 180° symmetry became more evident. We were then particularly interested in taking field sweeps at orientations 180° apart. By taking sums and differences of the data from two such orientations we could study the field dependence of both the odd and even field components. Such data were taken for the thermoelectric power at a mean temperature of 1.47°K. The even field component showed the oscillatory behavior as a function of H, while the odd field component did not. It is clear that further work is needed to explain the nature of the even and odd components.

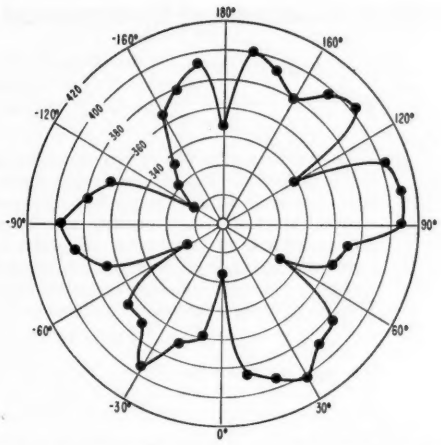


FIG. 8. Polar plot of thermoelectric voltage (microvolts) as a function of magnet angle for a bismuth single crystal. $H=12\,600$ gauss; the mean temperature=5.298°K and the temperature difference was 0.128°K. H was perpendicular to a binary axis when the magnet angle was 0°. H was perpendicular to the trigonal axis throughout the rotation.

As a result of the above longitudinal thermomagnetic measurements, we feel that the transverse thermomagnetic (Nernst and Righi-Leduc) effects in bismuth will exhibit an oscillatory magnetic dependence. One of us (J.B.) plans to make such transverse measurements. It is probable that other metals showing a de Haas–van Alphen effect will also exhibit oscillatory thermomagnetic dependencies. The thermomagnetic effects of such metals as graphite, beryllium, and zinc[23] are the most likely ones to examine since they exhibit large de Haas–van Alphen effects.

The authors wish to thank the members of the Cryogenics Branch for their helpful discussions.

[19] E. Grüneisen and J. Gielessen, Ann. Physik **26**, 449 (1936).
[20] E. Grüneisen and J. Gielessen, Ann. Physik **27**, 243 (1936).
[21] Grüneisen, Rausch, and Weiss, Ann. Physik **7**, 1 (1950).
[22] K. Rausch, Ann. Physik **1**, 190 (1947).
[23] P. B. Alers, Bull. Am. Phys. Soc. **30**, No. 1, 41 (1955).

Cyclotron Resonance of Electrons and Holes in Silicon and Germanium Crystals

G. DRESSELHAUS, A. F. KIP, AND C. KITTEL
Department of Physics, University of California, Berkeley, California

(Received December 16, 1954)

An experimental and theoretical discussion is given of the results of cyclotron resonance experiments on charge carriers in silicon and germanium single crystals near 4°K. A description is given of the light-modulation technique which gives good signal-to-noise ratios. Experiments with circularly polarized microwave radiation are described. A complete study of anisotropy effects is reported. The electron energy surfaces in germanium near the band edge are prolate spheroids oriented along ⟨111⟩ axes with longitudinal mass parameter $m_l=(1.58\pm0.04)m$ and transverse mass parameter $m_t=(0.082\pm0.001)m$. The electron energy surfaces in silicon are prolate spheroids oriented along ⟨100⟩ axes with $m_l=(0.97\pm0.02)m$; $m_t=(0.19\pm0.01)m$. The energy surfaces for holes in both germanium and silicon have the form

$$E(k)=Ak^2\pm[B^2k^4+C^2(k_x^2k_y^2+k_y^2k_z^2+k_z^2k_x^2)]^{\frac{1}{2}}.$$

We find, for germanium, $A=-(13.0\pm0.2)(\hbar^2/2m)$, $|B|=(8.9\pm0.1)(\hbar^2/2m)$, $|C|=(10.3\pm0.2)(\hbar^2/2m)$; and for silicon, $A=-(4.1\pm0.2)(\hbar^2/2m)$, $|B|=(1.6\pm0.2)(\hbar^2/2m)$, $|C|=(3.3\pm0.5)(\hbar^2/2m)$. A discussion of possible systematic errors in these constants is given in the paper.

1. INTRODUCTION

IN cyclotron resonance the current carriers in a solid are accelerated in spiral orbits about the axis of a static magnetic field H. The angular rotation frequency is

$$\omega_c=\pm eH/m^*c, \tag{1}$$

where m^* is the effective mass of the carrier. The experiment determines the effective mass directly, and is the first experiment to do so. Resonant absorption of energy from an rf electric field perpendicular to the static magnetic field occurs when the frequency of the rf field is equal to the cyclotron frequency $f_c=\omega_c/2\pi$. The motion is not unlike that of the particles in a cyclotron or simple magnetron. The $\pm$ choice in Eq. (1) indicates that holes and electrons will rotate in opposite senses in the magnetic field.

We consider the order of magnitude of several physical quantities relevant to the experiment. We make the estimates using $m^*/m\cong0.1$, which is not unrepresentative. For $f_c=24\,000$ Mc/sec, or $\omega_c=1.5\times10^{11}$ radians/sec, we have $H\cong860$ oersteds. The radius of the orbit is $r=v/\omega_c$. The mean radius for carriers in a Maxwellian velocity distribution at temperature T is

$$\bar{r}=\left(\frac{8kT}{\pi m^*}\right)^{\frac{1}{2}}\frac{1}{\omega_c}, \tag{2}$$

as $\bar{v}=(8kT/\pi m^*)^{\frac{1}{2}}$. For $T=4$°K, $\bar{v}\cong4\times10^6$ cm/sec, and $\bar{r}\cong3\times10^{-5}$ cm. The transition probability in cyclotron resonance is proportional to the square of the electric dipole moment; in electron spin resonance the transition probability is proportional to the square of the magnetic moment. As the maximum electric field in a resonant cavity is of the same order of magnitude as the maximum magnetic field, the ratio of the transition probabilities for cyclotron and for spin resonance will be of the order of

$$P_c/P_s\approx(e\bar{r})^2/\mu_B^2\approx10^{12} \tag{3}$$

for equal line widths, frequencies, and numbers of effective carriers. The Boltzmann factor in spin resonance is to be taken into account in the definition of an effective carrier. The substantial advantage favoring the detection of cyclotron resonance is partly lost because of the low carrier concentrations used in cyclotron resonance.

The line width is determined by the collision relaxation time τ, which describes the effect of collisions of the carriers with lattice vibrations, impurity atoms, and other imperfections. It is necessary that $\omega_c\tau\gtrsim1$ in order to obtain a distinctive resonance. In other words, the mean free path must be large enough so that the average carrier will get $1/2\pi$ of the way around a circle between successive collisions. For $\omega_c=1.5\times10^{11}$ sec^{-1}, we require $\tau\approx10^{-11}$ sec or longer. At room temperature the relaxation times of carriers in semiconductors and metals are commonly of the order of 10^{-13} to 10^{-15} second. It is usually necessary to work with high-purity crystals in the liquid hydrogen or liquid helium temperature range in order to obtain relaxation times which are long enough to permit the observation of cyclotron resonance with X- or K-band microwave equipment.

The theory of cyclotron resonance absorption is elementary, and for free particles goes back to Drude, Voigt, and Lorentz. Cyclotron resonance of free electrons in the earth's magnetic field has been observed in the propagation of radio waves in the ionosphere. The idea that it might be possible to carry out cyclotron resonance experiments in solids has been considered independently by a number of workers. In 1951 Dorfmann[1] published the suggestion of the possible application of cyclotron resonance to solids. Independently and simultaneously, Dingle[2] published his work on the

[1] J. Dorfmann, Doklady Acad. Sci. (U.S.S.R.) **81**, 765 (1951).
[2] R. B. Dingle, Ph.D. thesis, Cambridge University, 1951 (unpublished); Proceedings of the International Conference on Very Low Temperatures, edited by R. Bowers (Oxford, England, August 1951), p. 165; Proc. Roy. Soc. (London) **A212**, 38 (1952).

quantum theory of cyclotron resonance of a free particle, and also discussed the possible application to solids. Shockley[3] gave the solution of the problem of the cyclotron frequency for an ellipsoidal energy surface; his result is applicable directly to the conduction bands of silicon and germanium. He also derived expressions for the effective mobility in transverse and longitudinal cyclotron resonance. Later, Suhl and Pearson[4] reported an unsuccessful experimental attempt to detect cyclotron resonance in germanium at 77°K. The present authors[5] reported the first successful cyclotron resonance experiments, on germanium at 4°K. Our original results on germanium were incomplete, and important further developments for germanium have been reported by Lax, Zeiger, Dexter, and Rosenblum,[6] and by Dexter, Zeiger, and Lax.[7] The first work on silicon was reported concurrently by the Lincoln and Berkeley groups.

2. CLASSICAL THEORY OF CYCLOTRON RESONANCE FOR AN ISOTROPIC EFFECTIVE MASS

We give now a brief classical discussion of cyclotron resonance absorption by a carrier of isotropic effective mass. The theory will be generalized in later sections following a discussion of the experimental results. We review briefly the elementary classical theory of the process, assuming an isotropic effective mass m^* and an isotropic relaxation time τ, both independent of the velocity. In unpublished work we have developed the theory from the viewpoint of the Boltzmann transport equation,[8] but it is not worth while to reproduce the calculations here. The machinery of the Boltzmann equation is useful if one wishes to include a specific velocity dependence of the relaxation time, but we have no direct knowledge of the velocity dependence of the relaxation time in the circumstances of our experiments.

The equation of motion for the drift velocity is

$$m^*\left(\frac{d\mathbf{v}}{dt}+\frac{1}{\tau}\mathbf{v}\right)=e\left(\mathbf{E}+\frac{\mathbf{v}\times\mathbf{H}}{c}\right). \tag{4}$$

We take H as the static field along the z-axis and neglect the rf magnetic field. For plane-polarized radiation E_x, we have

$$m^*\left(i\omega+\frac{1}{\tau}\right)v_x=eE_x+\frac{e}{c}v_yH; \tag{5}$$

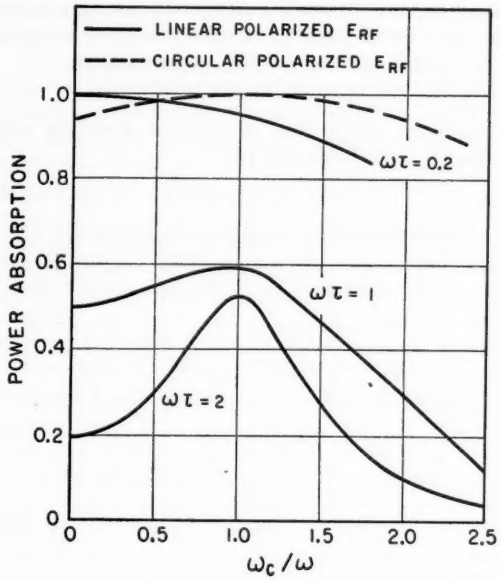


FIG. 1. Theoretical curves showing relative power absorption at constant frequency as a function of the static magnetic field intensity in units ω_c/ω, for various relaxation times in units $\omega\tau$. Curves are given for both linear polarization (Eq. [8]) and circular polarization of the rf field.

$$m^*\left(i\omega+\frac{1}{\tau}\right)v_y=-\frac{e}{c}v_xH.$$

We solve for v_x, finding for the complex conductivity,

$$\sigma=j_x/E_x=Nev_x/E_x=\sigma_0\left[\frac{1+i\omega\tau}{1+(\omega_c^2-\omega^2)\tau^2+2i\omega\tau}\right], \tag{6}$$

where

$$\sigma_0=Ne^2\tau/m^* \tag{7}$$

is the static conductivity; N is the carrier concentration. The losses are proportional to the real part of the conductivity. We express the result in convenient dimensionless form by writing $\nu=\omega\tau$, $\nu_c=\omega_c\tau$; the real part σ_R of σ is given by

$$\sigma_R/\sigma_0=\frac{1+\nu^2+\nu_c^2}{(1+\nu_c^2-\nu^2)^2+4\nu^2}. \tag{8}$$

This function is plotted in Fig. 1 for $\nu=0.2$, 1, and 2; it is seen that the resonance is quite well defined for $\nu=2$.

It is interesting to state in terms of the mobility the condition $\omega\tau>1$ for the observation of cyclotron resonance. In esu, $\tau=m^*\mu/e$; to have $\omega\tau>1$ requires $\mu>e/\omega m^*$. For $f=24\,000$ Mc/sec, the condition on μ expressed in practical units is, approximately,

$$(m^*/m)\mu>11\,000\ \text{cm}^2/\text{volt-sec}. \tag{9}$$

A high mobility does not in itself assure that a cyclotron resonance experiment is feasible; an appropriate average effective mass must also be considered. For narrow

[3] W. Shockley, Phys. Rev. **90**, 491 (1953).
[4] H. Suhl and G. L. Pearson, Phys. Rev. **92**, 858 (1953).
[5] Dresselhaus, Kip, and Kittel, Phys. Rev. **92**, 827 (1953).
[6] Lax, Zeiger, Dexter, and Rosenblum, Phys. Rev. **93**, 1418 (1954).
[7] Dexter, Zeiger, and Lax, Phys. Rev. **95**, 557 (1954).
[8] See, for example, the related calculation by R. Jancel and T. Kahan, J. phys. et radium **14**, 533 (1953); L. G. H. Huxley, Proc. Phys. Soc. (London) **B64**, 844 (1951).

energy gaps, the mobility divided by the gap energy may be a useful guide to relative relaxation times, as under some conditions $m^* \propto E_g$, approximately.

Several limiting situations are of interest:

(a) $\nu_c \gg \nu$; $\nu_c \gg 1$. This situation is found in very strong magnetic fields. We have

$$\sigma_R = \sigma_0/\nu_c^2 \propto 1/H^2, \tag{10}$$

so the losses well above the resonance field fall off as $1/H^2$. The carrier orbits are tightly coiled and very little drift is permitted in the direction of the electric field.

(b) $1 \gg \nu_c \gg \nu$. This situation is found at low frequencies at room temperature, or at low frequencies in weak magnetic fields at low temperatures. We have

$$\sigma_R/\sigma_0 \cong 1-\nu_c^2, \tag{11}$$

corresponding to a fractional resistivity change

$$\Delta\rho/\rho \cong \nu_c^2 = (\omega_c\tau)^2 = (\mu H/c)^2, \tag{12}$$

where the mobility is written as

$$\mu = e\tau/m^*. \tag{13}$$

This limit represents the low-frequency transverse magnetoresistance in the absence of a Hall electric field; in many actual problems, part of the transverse magnetoresistance is canceled[9] by the effect of the Hall field.

(c) $\nu \gg 1$; $\nu_c = 0$. This situation occurs in the infrared:

$$\sigma_R/\sigma_0 \cong 1/\nu^2. \tag{14}$$

(d) $\nu = \nu_c \gg 1$. This is the condition for cyclotron resonance. We have, from Eq. (8),

$$\sigma_R/\sigma_0 = 1/2. \tag{15}$$

Thus, at cyclotron resonance, the conductivity is one-half of the dc conductivity. If we had taken circularly polarized radiation in place of plane-polarized radiation, the ratio would have been unity. The factor one-half represents the selective absorption of one of the two circular components of a plane wave; the other component passes freely in the limit considered. The component which is absorbed at cyclotron resonance remains in phase with the drift velocity throughout the motion, just as in ordinary dc conductivity; hence the absorption of this component is identical with the dc absorption.

The rf conductivity at resonance is related to the rf conductivity at zero magnetic field by the ratio

$$\sigma_R(\text{res})/\sigma_R(H=0) = \nu_c^2/2, \tag{16}$$

provided $\nu_c \gg 1$. For $\omega_c\tau \approx 10$, the ratio $\sigma_R(\text{res})/\sigma_R(H=0)$ is of the order of 50.

We now consider the Q of the sample at cyclotron resonance. We suppose the specimen is located in the microwave cavity at a position of negligible rf magnetic field. The stored energy density is $\epsilon\langle E^2\rangle_{Av}/8\pi$; the energy dissipated per radian at resonance is $\sigma_R\langle E^2\rangle_{Av}/\omega$, so

$$Q_{\text{cycl}} = \epsilon\omega/8\pi\sigma_R = m^*\epsilon\omega/4\pi Ne^2\tau. \tag{17}$$

For the standard example described in the introduction,

$$Q_{\text{cycl}} \approx 10^{12}/N. \tag{18}$$

This may be compared with the estimated Q for electron spin resonance[10]:

$$Q_{\text{spin}} = (kT/N\mu_B^2)(\Delta\omega/\omega). \tag{19}$$

For the sake of comparison, we set the spin resonance line width equal to the cyclotron resonance line width. We take $\Delta\omega = 1/\tau$, so

$$Q_{\text{spin}} \cong kT/N\mu_B^2\omega\tau \approx 10^{24}/N \tag{20}$$

under the previous conditions. The requirement that a resonance may be detected may be stated roughly that for a given line width a certain threshold $1/Q$ must be exceeded, $1/Q$ being the measure of the absorption. It appears therefore that cyclotron resonance should be detectable at carrier concentrations lower by a factor of the order of 10^{12} than spin resonance; the factor arises principally from the relevant matrix elements, as we saw earlier in Eq. (3). There are several practical considerations which act to reduce the factor, but it would appear that it is in principle within the range of existing equipment to detect 10^4 carriers in cyclotron resonance.

The half-width at half σ_R on the σ_R *vs* ω curve is determined by the condition

$$\tau\Delta\omega = 1. \tag{21}$$

For broad lines the position of maximum absorption shifts slightly toward higher H. For $\omega\tau = 1$ the fractional shift is approximately 1/9; for $\omega\tau \gg 1$ the fractional shift is of the order of $1/8(\omega\tau)^2$, which will usually be unimportant. In our experiments on germanium and silicon at 4°K and 24 000 Mc/sec, the value of $\omega\tau$ was about 10.

Depolarization Effects

We show now that the electrostatic self-interaction of the resonance polarization may be neglected at the lower carrier concentrations with which we are concerned, although at higher concentrations new and undesirable effects enter. We consider the depolarizing fields associated with the shape of the sample; the effect of possible Lorentz fields is neglected. We suppose for simplicity that the sample has axial symmetry about the axis of the static magnetic field. In the axial plane the internal electric field is

$$E_i = E - LP, \tag{22}$$

where L is the depolarizing factor. The polarization P

[9] A. H. Wilson, *Theory of Metals* (Cambridge University Press, London, 1953), second edition p. 215.

[10] C. Kittel and J. M. Luttinger, Phys. Rev. **73**, 162 (1948).

is given by

$$P=\chi_0 E_i+Ne\int v dt=\chi_0 E_i-iNev/\omega; \quad (23)$$

here $\chi_0=(\epsilon-1)/4\pi$, and ϵ is the dielectric constant of the host crystal exclusive of carriers. The internal field is

$$E_i=\frac{E+i(LNev/\omega)}{1+L\chi_0}, \quad (24)$$

and the equation of motion becomes ($L_i=L/(1+L\chi_0)$).

$$m^*\left[i\left(\omega-\frac{L_iNe^2}{m^*\omega}\right)+\frac{1}{\tau}\right]\mathbf{v}=\frac{e}{1+L\chi_0}\mathbf{E}+(e/c)\mathbf{v}\times\mathbf{H}. \quad (25)$$

The effect of the carrier polarization is to replace ω in the equation for v by $\omega[1-(\omega_p/\omega)^2]$, where the plasma frequency ω_p is given by

$$\omega_p=(L_iNe^2/m^*)^{\frac{1}{2}}=(L_i\sigma_0/\tau)^{\frac{1}{2}}. \quad (26)$$

The effect will be important when ω_p is of the order of ω or larger. A critical concentration N_p may be defined by the relation

$$N_p=m^*\omega^2/L_ie^2. \quad (27)$$

In the standard example $N_p\cong 4\times 10^{12}/L$ cm^{-3}. For a sphere $L=4\pi/3$, so $N_p\cong 10^{12}$ cm^{-3}; for a flat disk one could quite easily get N_p up to 10^{14} cm^{-3}. In germanium with 10^{14} impurity atoms/cm^3 it appears that depolarization effects will enter somewhat above 10°K. It is very important to avoid depolarization effects, and this may be done by the use of thin specimens at low concentrations, with the rf electric field parallel to the plane of the specimen and the static magnetic field normal to the plane.

Depolarization effects can produce a fictitious cyclotron resonance in a limited temperature range in the following way: If $\omega\tau$ is too small for the normal cyclotron resonance to be observable, it may still be possible to have $\omega_p^2\tau/\omega\gtrsim 1$, so that a magnetoplasma resonance will appear at a field H such that

$$\omega\omega_c=\omega^2-\omega_p^2, \quad (28)$$

where ω_c is the cyclotron angular frequency. This equation describes the Zeeman effect of an oscillator with the natural frequency ω_p. If ω^2 may be neglected, we have the resonance condition

$$-\omega\omega_c\cong\omega_p^2=L_i\sigma_0/\tau=L_iNe^2/m^*. \quad (29)$$

We have observed in p-Ge near 20°K a resonance which possibly may arise from this effect; unlike the actual cyclotron resonances, the resonance field varied strongly with temperature and also with the shape of the specimen. A displacement of the cyclotron resonances in silicon has also been observed under conditions of higher carrier concentrations. A distinctive magnetoplasma resonance line has been observed in n-InSb; the line showed the expected dependence on frequency and on sample shape.

The magnetoplasma effect[11] appears to impose an upper limit to the carrier concentrations at which cyclotron resonance may be observed. It is hoped that the effect will be investigated further in order to establish whether or not the predicted limitation actually occurs. A background of nonresonant carriers is not necessarily troublesome, so it may be possible, for example, to detect cyclotron resonance of a subcritical concentration of electrons in the presence of a supercritical concentration of holes. A further apparent difficulty which may enter at the high carrier concentrations encountered in metals is that the diameter of the cyclotron orbit may be large in comparison with the skin depth. Unfortunately the relevant transport problem has not yet been solved, so one does not know to what extent the line shape and intensity depend on the ratio of orbit diameter to skin depth. In superconductors one has the plasma effect, and also the difficulty in obtaining penetration of the static magnetic field. Even if a superconducting film is used which is thin in comparison with the penetration depth for parallel static magnetic fields, it is not possible[12] for a field *normal* to the surface to penetrate uniformly over a useful area. A preliminary attempt in this laboratory by G. Feher to detect cyclotron resonance in a thin superconducting film was not successful; the negative result is not surprising, in view of the foregoing objections.

3. EXPERIMENTAL METHODS

A brief account will be given of the microwave apparatus which has been used in these experiments. The experimental techniques will be described, including the various methods which have been used to ionize electrons and holes. Sample preparation and mounting arrangements will be discussed. Some experiments which allowed discrimination between resonance absorption due to electrons and holes will be described.

Apparatus

In order to fulfill the requirement that $\omega_c\tau\geq 1$, the experiments have been performed at microwave frequencies, at or near 4°K, in the range of 9000 and 24 000 Mc/sec. The apparatus used is essentially the same as for conventional paramagnetic resonance experiments, except that the geometry is modified so that the microwave electric field is perpendicular to the external applied magnetic field. This is in contrast to the paramagnetic arrangement which puts the micro-

[11] C. Kittel, Proceedings of the 10th Solvay Congress, 1954 (unpublished).

[12] We consider the maximum radius R which a thin superconducting disk of thickness $a\ll$penetration depth d may have without significant perturbation of the normal field H_0. By the London equation $\mathbf{j}=\mathbf{r}\times\mathbf{H}_0/2\Lambda c$, this current produces in turn a magnetic field $\Delta H=(2\pi a/c)\int(j/r)dr=(2\pi\Lambda c^2)aRH_0$. We will have $(\Delta H/H_0)\ll 1$ if $aR\ll 2d^2$, using the relation $d^2=\Lambda c^2/4\pi$. This effect was suggested by G. Feher.

wave magnetic field perpendicular to the external magnetic field. Absorption of energy under cyclotron resonance conditions has been determined by measuring the change in the Q of a microwave cavity in which the sample has been placed. The applied magnetic field may be varied in order to obtain the spectrum of power absorption in the sample *vs* the magnetic field. From the absorption spectrum the effective masses may be obtained directly from the equations derived in Secs. 4 and 5 below.

The microwave circuit is a conventional one. Microwave power from a stabilized klystron is fed into the test cavity through a magic tee; 2K39 and 2K33 klystrons have been used at the lower and higher frequencies, respectively. Some of the power reflected back from the test cavity reaches a crystal detector in another arm of the tee. The fourth arm of the tee is fitted with a matched load so that all power incident on the crystal has been reflected from the cavity. A change in the loaded Q of the cavity resulting from cyclotron absorption in the sample causes a change in the power reflected to the crystal. This change in reflected power is proportional to the loss in the sample for small signals (i.e., for sample losses small compared to other losses in the cavity). For small signals, the change in output voltage of the crystal will be proportional to the change in power incident upon it. Thus, the crystal gives a voltage signal which measures the variation of power absorbed by the sample as the magnetic field is varied.

One of the several modulation techniques is used to produce an ac signal at the crystal. The ac signal is passed through a detection channel consisting of an amplifier and a lock-in detector, the output of which is put on a pen recorder. A 1000-cps modulation frequency is used. The external magnetic field is caused to sweep slowly from zero to a maximum of 10 000 oersteds by means of an electronic control on the field coil of a motor generator used to provide the magnet current. Field markers are applied periodically to the pen recorder. These markers are obtained from a rotating coil in the magnetic field. The signal from this coil is balanced on a potentiometer against another signal from a coil mounted on the same shaft which rotates in the field of a fixed permanent magnet. The value of any given field is determined by the potentiometer setting which gives a null signal.

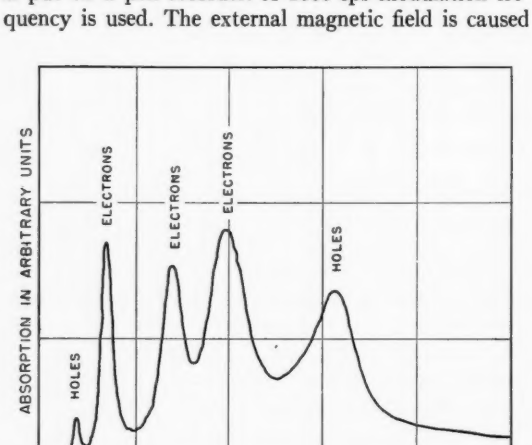


FIG. 2. Typical cyclotron resonance results in germanium near 24 000 Mc/sec and 4°K: direct copy from a recorder trace of power absorption *vs* static magnetic field in an orientation in a (110) plane at 60° from a [100] axis.

In all the experiments low temperatures are provided by liquid helium. In most experiments liquid helium is allowed to enter the cavity, so that the sample is immersed in helium. In some experiments pumping on the helium has allowed reduction of the temperature to 2°K.

In experiments with plane-polarized microwaves the samples were placed inside rectangular cavities made from wave guide stock. The cavity was coupled to the wave guide through an iris containing an appropriate coupling hole. The construction of cavities used in the circularly polarized microwave experiments will be discussed below under that heading.

Experimental Technique

At the low temperatures required in these experiments, the equilibrium number of free charge carriers is usually too small to allow observation of resonance absorption. In the original experiments on germanium it was found that in both frequency ranges used the microwave electric field in the cavity was sufficient to cause ionization of impurity atoms by multiplication processes taking place in the sample. Only electrons or only holes appear to be produced by this ionization process, depending on whether n- or p-type germanium is used. This results from the fact that energies given to the carriers by the microwave field are only enough to cause ionization of impurity atoms ($\sim$0.01 ev in germanium), and not enough to produce electron-hole pairs by the removal of electrons from the valence band to the conduction band (0.7 ev in germanium). In silicon, where the ionization energy of the impurity atoms is $\sim$0.05 ev, the available microwave power is insufficient to produce ionization, and other methods must therefore be used to produce free charge carriers. In experiments using the microwave ionization technique, the microwave power was amplitude-modulated in order to provide an ac signal for the detection channel.

The microwave ionization method provides fairly satisfactory information on the positions of peaks in the absorption curve. However, the observed widths of the resonance lines may vary between wide limits, depending on the microwave power level used. This effect results from the dependence of the multiplication process on the applied magnetic field. Thus, near a resonance peak, charge carriers pick up more energy from the microwave electric field and hence cause more

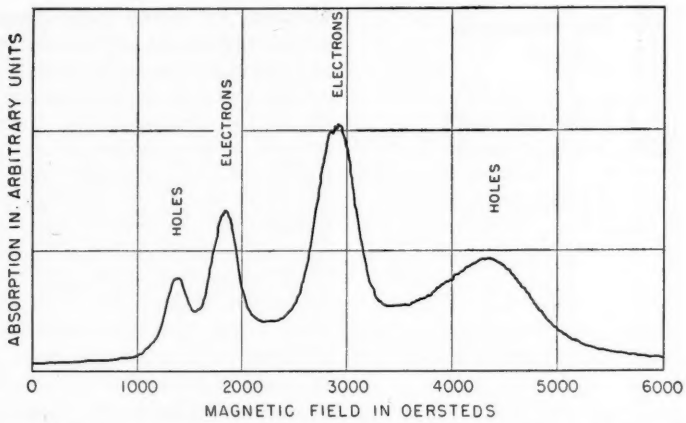


FIG. 3. Typical cyclotron resonance results in silicon near 24 000 Mc/sec and 4°K: static magnetic field orientation in a (110) plane at 30° from a [100] axis.

multiplication. Since the size of the absorption peak will vary with the number of carriers present, peak heights can be enormously increased, resulting in apparently very narrow lines. Lines as narrow as 10 oersteds have been observed in some cases where microwave power levels were so low that effective multiplication occurred only very close to the resonance peak.

A more satisfactory method of producing free charge carriers involves the use of light excitation; it was first used by Dexter, Zeiger, and Lax.[13] In this technique, light from an incandescent source is focused on the sample through a hole in the cavity. Since the light produces hole-electron pairs, both hole and electron resonance is observed, regardless of whether the sample is *n*- or *p*-type. The first observations on silicon were made by this technique.[14] In our first experiments using light excitation, amplitude modulation of the microwaves was used to provide an ac signal.

A slight modification of the optical excitation method is the most satisfactory method so far used.[15] The modification consists in modulating the light beam by means of a rotating disk. This results in a modulation of the free carrier density, which in turn gives a modulated microwave absorption signal for operation of the detection channel. An auxiliary light source incident on a photocell is modulated by the same rotating disk. After amplification the signal from the photocell is used as the reference signal for the lock-in detector. The light-modulation method gives a very large improvement in signal-to-noise ratio over the earlier techniques. Figures 2 and 3 show typical recorder tracings of resonance lines in germanium and silicon crystals.

Because of the high dielectric constants of both silicon and germanium, samples placed in a cavity seriously perturb the resonant frequency of the cavity. Furthermore, since at all but very low temperatures the samples are very lossy because of their high conductivity, it is difficult to determine the resonant frequency without cooling to liquid-helium temperature. Both problems are minimized by using small samples. A typical sample is a disk about 3 mm in diameter and about 0.5 mm thick. Samples have usually been prepared by rough-cutting from a single crystal, grinding to size with abrasives, and etching the surface for several minutes in an etch made up of 1 cc HF, 1 cc H_2O_2 (30 percent), and 4 cc H_2O.

The importance of anisotropy in the effective masses of silicon and germanium requires that data be obtained as a function of crystal orientation. A sample cut with its surface in the (110) plane, oriented so that the applied field can be directed along the [001], [1$\bar{1}$0], and [1$\bar{1}$1] directions by rotation of the sample, provides all of the necessary anisotropy information. To allow this rotation while the cavity is immersed in liquid helium, the sample is fastened with coil dope to a mushroom-shaped Lucite holder. This holder is put inside the cavity, with its stem inserted through a small hole in the broad face of the cavity into a hole in the middle of a brass wheel outside the cavity. The wheel can be rotated from outside the Dewar in 1° steps. The light beam shines on the sample through another hole in the opposite side of the cavity. The samples were oriented by x-ray diffraction measurements by Professor J. Washburn. The effective mass data are presented in Secs. 4 and 5 below.

In the original experiments on germanium it was of interest to verify the sign of the charge carriers involved. For this reason, experiments using circularly polarized microwaves were carried out, using *n*- and *p*-type samples in both frequency ranges. Absorption was observed only when the direction of circular polarization corresponded to the direction of rotation of the charge carriers.

[13] Dexter, Zeiger, and Lax, Phys. Rev. **95**, 557 (1954).
[14] Dexter, Lax, Kip, and Dresselhaus, Phys. Rev. **96**, 222 (1954).
[15] A. F. Kip, Physica **20**, 813 (1954).

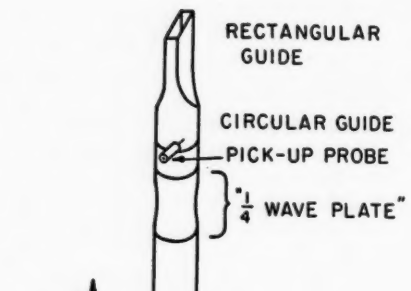


FIG. 4. Experimental arrangement for circular polarization studies of cyclotron resonance.

A description will be given of the problems involved in the production and use of circularly polarized microwaves. The klystron power was taken from a conventional rectangular guide through a gradual transition into circular guide, as shown in Fig. 4. The microwaves then passed through a microwave quarter-wave plate into another guide of circular cross section, and thence through a circular iris into a cylindrical cavity. The quarter-wave plate consisted simply of a section of several wavelengths of circular guide which had been squeezed into an appropriate elliptical shape. If the polarized microwaves from the rectangular guide are passed into this elliptical section with the plane of polarization at 45° to the axes of the ellipse, two mutually perpendicular modes of equal intensity will be transmitted through the section. Since the wave velocity depends on the transverse guide dimension perpendicular to the E vector, these two modes will travel with different velocities. Adjustment of ellipticity and length of the elliptical section produces a quarter wavelength shift in the phase of the two modes. This adjustment is made empirically by use of an analyzer placed at the position of the cavity at the end of the circular guide. The analyzer is constructed of a circular guide which makes a transition to a rectangular guide. A crystal detector is placed at the termination of the rectangular guide. Since the rectangular guide will transmit only a polarized wave, the analyzer will measure the intensity of the wave polarized in any given plane, depending on the analyzer orientation. When complete circular polarization is achieved, the power picked up by the crystal is independent of the orientation of the analyzer. In order to prevent reflection of power not accepted by the rectangular guide, both the analyzer and the transition from rectangular to circular guide are provided with absorbing fins placed in the circular guide. These fins are so oriented as to absorb all microwave components not transmitted or accepted by the rectangular guide.

Once circularly polarized waves are incident on the iris of the circular cavity, there remains only the problem of insuring that the cavity has perfect microwave cylindrical symmetry. Any departure from this symmetry will result in two different resonance frequencies for the two mutually perpendicular modes into which the circularly polarized mode can be analyzed. In general, satisfactory symmetry will not be automatically achieved; therefore, provision must be made for adjustment of symmetry. This is done in the following way: A small rectangular fin of polystyrene is placed in the end of the cylindrical cavity. The fin can be rotated about the axis of the cavity, and can be inserted a variable distance from the end wall of the cavity. The cylindrical axis passes through the plane of the fin. The plane of the fin is rotated until it contains the E vector of the plane-polarized component for which the wavelength in the cavity is longest. Because of its dielectric constant the polystyrene makes the cavity look longer to this mode without materially affecting the perpendicular mode. Once proper orientation is obtained, the fin is inserted further into the cavity where the E field is higher and hence the perturbation is greater. The adjustment is made empirically until the two perpendicular modes are completely degenerate. The method involves sweeping the klystron frequency and adjusting the fin until the two resonant modes of the cavity are made to coincide in frequency. After adjustment of the cavity, the small cylindrical sample is placed accurately on the axis of the cavity, so as to maintain the degeneracy of the two modes. The sample is held in position by partially filling the cavity with layers of tightly fitting Styrofoam, between which the sample is placed.

The problem of detection of the power reflected from the cylindrical cavity is very simple with the arrangement used. A pickup probe is placed in the cylindrical guide of the klystron side of the quarter-wave plate. This probe is oriented perpendicularly to the E vector of the wave coming from the klystron through the rectangular guide, and hence does not pick up a signal from this wave. However, after the wave travels through the quarter-wave plate into the cavity and is reflected back through the quarter-wave plate again, it has been rotated through 90° and hence is picked up by the probe. The probe leads to a crystal detector. Amplitude modulation of the microwave power allows the usual detection channel to be used on the output signal from the crystal.

The magnetic field must be perpendicular to the microwave E vector, and hence must be along the axis of the cylindrical cavity. We therefore used a solenoidal air core electromagnet into which the Dewars and cavity were inserted. The direction of circular polarization obtained was confirmed by observing electron

spin resonance in an organic free radical, this resonance occurring in the same sense as electron cyclotron resonance. Circular polarization has been used only on germanium, where microwave ionization is possible.

A third method of ionization has also been used which allows the selective observation of one sign of charge carrier, depending on whether n- or p-type materials are used. In this method voltage is applied to the sample through soldered contacts. The voltage is modulated at the standard 1000-cps rate to give the required ac signal for the detection channel. This method gives the same selective production of conduction electrons or holes as given by the microwave ionization method. It has the advantage of applicability to silicon, and gives much better signal-to-noise ratio. The resonance peaks observed are broader than for optically excited carriers, probably because of shorter relaxation times involved. Shorter relaxation times are probably the result of the higher carrier energies produced by the applied voltage. In this method the sample is placed just outside a small hole in the broad face of the rectangular cavity. This allows the sample to be seen by the microwaves and at the same time avoids the problems involved in bringing the wires carrying the voltage into the cavity.

4. THEORY OF CYCLOTRON RESONANCE IN THE CONDUCTION BAND OF GERMANIUM AND SILICON

The neighborhood of the conduction band edge in both germanium and silicon consists of a set of spheroidal energy surfaces located in equivalent positions in k-space. We discuss now the theory of cyclotron resonance for surfaces of this character. We choose Cartesian coordinate axes with the z-axis parallel to the figure axis of the spheroid, and we measure the wave vector components from the center of the spheroid. For points in k-space sufficiently close to a band edge point, the energy is described by the equation

$$E(\mathbf{k})=\hbar^2\left(\frac{k_x^2+k_y^2}{2m_t}+\frac{k_z^2}{2m_l}\right). \tag{30}$$

Here m_l is the longitudinal mass parameter and m_t is the transverse mass parameter. We have no evidence as to the range in k-space over which this expression is adequate; no departures were observed in our experiments.

We wish now to discuss the energy levels in the presence of a uniform static field H. The usual procedure is to take the effective Hamiltonian,

$$\mathcal{H}(\mathbf{P})=\frac{P_x^2+P_y^2}{2m_t}+\frac{P_z^2}{2m_l}, \tag{31}$$

and solve the equations of motion

$$\mathbf{v}=\nabla_P\mathcal{H}(\mathbf{P}); \tag{32}$$

$$d\mathbf{P}/dt=e[\mathbf{E}+(1/c)\mathbf{v}\times\mathbf{H}]. \tag{33}$$

Here $\mathbf{P}=\mathbf{p}-e\mathbf{A}/c$, where p is the momentum and $\mathbf{A}$ the vector potential.

This procedure in a restricted form was discussed by Jones and Zener[16]; recent discussions have been given by Shockley,[17] Luttinger,[18] and Adams.[19]

Shockley[20] has given the solution of the cyclotron frequency problem for a general ellipsoidal energy surface. We indicate the method of solution here. For the spheroidal surface (31),

$$\mathbf{v}=(P_x/m_t;\ P_y/m_t;\ P_z/m_l). \tag{34}$$

We take

$$\mathbf{H}=H(\sin\theta;\ 0;\ \cos\theta). \tag{35}$$

Then Eq. (33) becomes, letting $\omega_t=eH/m_tc$ and $\omega_l=eH/m_lc$,

$$\begin{aligned} i\omega P_x-\omega_l P_y\cos\theta&=0;\\ i\omega P_y-\omega_l P_z\sin\theta+\omega_t P_x\cos\theta&=0;\\ i\omega P_z+\omega_t P_y\sin\theta&=0. \end{aligned} \tag{36}$$

The associated secular equation has the solution

$$\omega^2=\omega_t^2\cos^2\theta+\omega_t\omega_l\sin^2\theta. \tag{37}$$

Thus, the effective mass determining the cyclotron frequency when the static magnetic field makes an angle θ with the longitudinal axis of the energy surface is

$$\left(\frac{1}{m^*}\right)^2=\frac{\cos^2\theta}{m_t^2}+\frac{\sin^2\theta}{m_tm_l}. \tag{38}$$

In Fig. 5 we give a plot of the experimental points obtained for electrons in germanium at 4°K as a function of the angle between the direction of the static magnetic field in a (110) plane and a [001] direction lying in the plane. The mass values derived from the theoretical fit to the experimental points are $m_l=(1.58\pm0.04)m$ and $m_t=(0.082\pm0.001)m$; we assume that there are a set of crystallographically equivalent energy spheroids oriented along the ⟨111⟩ directions in the Brillouin zone. Lax[21] *et al.* have reported $m_l=1.3m$ and $m_t=0.08m$ from a similar experiment. Our original observation[22] of one line in the [100] direction with $m^*=0.11m$ is in fair agreement with the later results.

In Fig. 6 we give a plot of the experimental points

[16] H. Jones and C. Zener, Proc. Roy. Soc. (London) **A144**, 101 (1934).

[17] W. Shockley, *Electrons and Holes in Semiconductors* (D. van Nostrand Company, New York, 1950), pp. 424 ff.

[18] J. M. Luttinger, Phys. Rev. **84**, 814 (1951); see also J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955).

[19] E. N. Adams, Phys. Rev. **85**, 41 (1952); **89**, 633 (1953).

[20] W. Shockley, Phys. Rev. **90**, 491 (1953); the problem had arisen also in connection with the de Haas-van Alphen effect.

[21] Lax, Zeiger, Dexter, and Rosenblum, Phys. Rev. **93**, 1418 (1954).

[22] See reference 4.

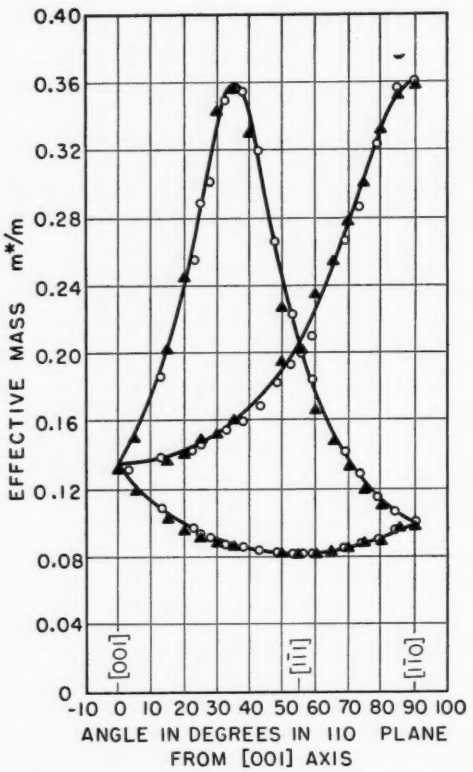


FIG. 5. Effective mass of electrons in germanium at 4°K for magnetic field directions in a (110) plane; the theoretical curves are calculated from Eq. (38), with $m_l=1.58m$; $m_t=0.082m$. The different types of points indicate different runs.

obtained for electrons in silicon at 4°K as a function of the angle between the direction of the static magnetic field in a (110) plane and a [001] direction lying in the plane. The theoretical curves are drawn for $m_l=(0.97\pm0.02)m$ and $m_t=(0.19\pm0.01)m$; we assume that there are a set of crystallographically equivalent energy spheroids oriented along the ⟨100⟩ directions in the Brillouin zone. In earlier work[23] the values $m_l=0.98m$ and $m_t=0.19m$ were reported under similar conditions.

Theoretical calculations of band structure have not reached as yet a state of development which permits the deductive derivation of the central features of the conduction band energy surfaces found experimentally. The most ambitious theoretical program has been that undertaken by F. Herman for germanium, but the band edge points turn out to be too sensitive to the details of the calculation to be reliable. Herman[24] has suggested, however, that the conduction band energy minima in silicon and germanium may arise from different bands: in silicon the band which at $k=0$ is a representation Γ_{15}^- of the cubic group is thought to lie lowest, whereas in germanium the lowest band at $k=0$ is thought to be a representation of Γ_2^- of the cubic group.

The structure of the conduction band edge of germanium determined by cyclotron resonance is consistent with the interpretation by Meiboom and Abeles[25] and by Shibuya[26] of magnetoresistance measurements on n-Ge by Estermann and Foner[27] and by Pearson and Suhl.[28] Similarly, the cyclotron resonance results for the conduction band edge of silicon are consistent with the magnetoresistance results of Pearson and Herring[29] on n-Si. In fact, the assignment of the energy surfaces in silicon to electrons or holes depends on the correlation with the magnetoresistance data.

In Sec. 3, it was found that in a circular polarization experiment on an n-Ge crystal in conditions of rf ionization with $H\|[100]$, absorption was observed only for one sense of the static magnetic field and not for the opposite sense. If the orbit of an electron is circular, absorption of circularly polarized radiation should occur only for one sense, but with an elliptical orbit there should be some absorption also in the opposite sense of the static field. We now calculate this absorption for a general orientation of the energy surface relative to the static magnetic field. The equations of motion are, using (33), (35), (36) and including an isotropic relaxation time τ,

$$(i\omega+1/\tau)P_x-\omega_t P_y=eE; \tag{39}$$

$$(i\omega+1/\tau)P_y+(\omega_0^2/\omega_t)P_x=-ieE. \tag{40}$$

The coordinate axes are chosen with the static magnetic field in the z-direction and the rf electric field in the xy plane; the energy surface is rewritten so that the principal axis of the surface makes an angle θ with the z-axis. The solution is independent of P_z, and we have set $P_z=0$ for convenience. If $\omega\tau\gg1$, the ratio $\mathfrak{R}$ of the power absorption at resonance in the weak sense of rotation to that in the strong sense of rotation is found to be

$$\mathfrak{R}=\left(\frac{m^*-m_t}{m^*+m_t}\right)^2. \tag{41}$$

For electrons in germanium with the static magnetic field in the [100] direction we have $\mathfrak{R}\cong0.06$, which is not inconsistent with the observations.

5. THEORY OF CYCLOTRON RESONANCE IN THE VALENCE BAND

The structures of the valence band edges of germanium and silicon are qualitatively similar. We discuss first the theory of the form of the energy surfaces near the band edge and secondly, the connection between

[23] Dexter, Lax, Kip, and Dresselhaus, Phys. Rev. **96**, 222 (1954).
[24] F. Herman, Phys. Rev. **95**, 847 (1954).
[25] S. Meiboom and B. Abeles, Phys. Rev. **95**, 31 (1954).
[26] M. Shibuya, Phys. Rev. **95**, 1385 (1954).
[27] I. Estermann and A. Foner, Phys. Rev. **79**, 365 (1950).
[28] G. L. Pearson and H. Suhl, Phys. Rev. **83**, 768 (1951).
[29] G. L. Pearson and C. Herring, Physica **20**, 975 (1954).

the cyclotron frequencies and the parameters which define the energy surfaces. The complete details of the calculations will be given in the doctoral thesis of G. Dresselhaus, of which a limited number of copies may be available for distribution by request late in 1955.

Everything we know at present indicates that the valence band edge occurs at the center of the Brillouin zone ($k=0$), at which point the band edge state has a threefold orbital degeneracy if the spin-orbit interaction is not considered. According to calculations by Herman[30] and others, it is most likely that the degenerate wave functions transform under the operations of the full cubic group according to the representation Γ_{25}^+ in the notation of Bouckaert, Smoluchowski, and Wigner,[31] or Γ_5^+ in the rotation of Von der Lage and Bethe.[32] In chemical language, the degenerate wave functions at the valence band edge have the transformation properties of p-orbitals arranged with opposite sign (bonding) on each of the two fcc lattices which compose the diamond structure. With spin-orbit interaction we have to deal at the valence band edge with bonding $p_{\frac{3}{2}}$ orbitals. The treatment below is quite general within the scope of the one-electron approximation; we do not make a tight binding assumption.

In order to establish a notation we first set up the solution to the problem without spin-orbit interaction, as has been discussed briefly by Shockley.[33] We will then extend the treatment to the actual problem with spin-orbit interaction. We make use of the pseudo-Bloch function representation introduced by Kittel and Mitchell.[34]

We make an arbitrary choice of a basis for the representation Γ_{25}^+ at $k=0$, taking the three degenerate states to transform as $\epsilon_1^+\sim yz$; $\epsilon_2^+\sim zx$; $\epsilon_3^+\sim xy$, following here, as below, the notation of Von der Lage and Bethe. We now construct three pseudo-Bloch functions from the original basis; that is, we construct by perturbation theory three functions $u_{\mathbf{k}}{}^i(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}$ which are eigenfunctions of the crystal translation operator but which are not in general eigenfunctions to the first order in $\mathbf{k}$ of the Hamiltonian. However, linear combinations of the $u_{\mathbf{k}}{}^i$ diagonalize the Hamiltonian to the first order in $\mathbf{k}$. The perturbation term in the Hamiltonian is $\mathcal{H}'=\hbar\mathbf{k}\cdot\mathbf{p}/m$, where $\mathbf{p}=-i\hbar\nabla$ is the momentum operator. Thus,

$$u_{\mathbf{k}}{}^i(\mathbf{r})=\epsilon_i{}^+ + (\hbar/m)\mathbf{k}\cdot\sum_{l\alpha j}\frac{|l\alpha j\rangle\langle l\alpha j|\,\mathbf{p}\,|i+\rangle}{E_0-E_{l\alpha}}, \quad (42)$$

where $l\alpha j$ denotes the state j belonging to the representation α in the band l; E_l is the energy of the lth band at $\mathbf{k}=0$.

[30] F. Herman, Physica **20**, 801 (1954).
[31] Bouckaert, Smoluchowski, and Wigner, Phys. Rev. **50**, 58 (1936).
[32] F. C. Von der Lage and H. Bethe, Phys. Rev. **71**, 612 (1947).
[33] W. Shockley, Phys. Rev. **78**, 173 (1950).
[34] C. Kittel and A. H. Mitchell, Phys. Rev. **96**, 1488 (1954).

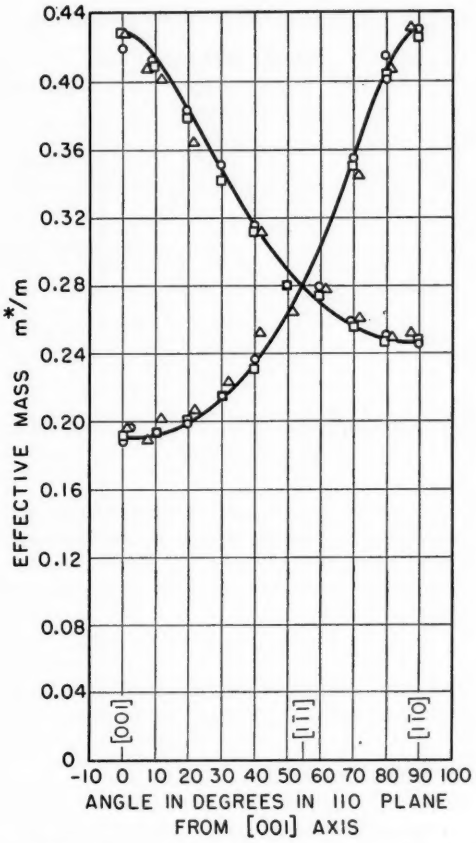


FIG. 6. Effective mass of electrons in silicon at 4°K for magnetic field directions in a (110) plane; the theoretical curves are calculated from Eq. (38), with $m_l=0.98m$; $m_t=0.19m$.

The perturbation matrix in this representation has the form

$$\langle r|\mathcal{H}'|s\rangle=\frac{\hbar^2}{m^2}\sum_{l\alpha j}\frac{\langle r+|\mathbf{k}\cdot\mathbf{p}|l\alpha j\rangle\langle l\alpha j|\mathbf{k}\cdot\mathbf{p}|s+\rangle}{E_0-E_{l\alpha}}, \quad (43)$$

as the matrix elements of $\mathbf{p}$ among the states $\epsilon_i{}^+$ are all zero. We can determine the dependence of $\langle r|\mathcal{H}'|s\rangle$ on the components of $\mathbf{k}$ by a simple observation. If all the $E_{l\alpha}$ were equal, say to E_1, the sum above could be carried out, giving

$$\langle r|\mathcal{H}'|s\rangle=\frac{\hbar^2}{m^2}\frac{\langle r|\mathcal{H}'^2|s\rangle}{E_0-E_1}. \quad (44)$$

We see immediately from the transformation properties of the $\epsilon_i{}^+$ that

$$\langle 1+|\mathcal{H}'^2|2+\rangle\propto k_xk_y, \quad (45)$$

with similar relations for other matrix elements. The form of each element as determined in this way will not change as we relax the above restriction on $E_{l\alpha}$.

The perturbation matrix is clearly of the form

$$\begin{vmatrix} Lk_x^2+M(k_y^2+k_z^2)-\lambda & Nk_xk_y & Nk_xk_z \\ Nk_xk_y & Lk_y^2+M(k_x^2+k_z^2)-\lambda & Nk_yk_z \\ Nk_xk_z & Nk_yk_z & Lk_z^2+M(k_x^2+k_y^2)-\lambda \end{vmatrix}=0. \tag{46}$$

Here

$$L=\frac{\hbar^2}{m^2}\sum_{l\alpha j}\frac{\langle 1+|p_x|l\alpha j\rangle\langle l\alpha j|p_x|1+\rangle}{E_0-E_{l\alpha}};$$

$$M=\frac{\hbar^2}{m^2}\sum_{l\alpha j}\frac{\langle 1+|p_y|l\alpha j\rangle\langle l\alpha j|p_y|1+\rangle}{E_0-E_{l\alpha}};$$

$$N=\frac{\hbar^2}{m^2}\sum_{l\alpha j}\frac{\langle 1+|p_x|l\alpha j\rangle\langle l\alpha j|p_y|2+\rangle+\langle 1+|p_y|l\alpha j\rangle\langle l\alpha j|p_x|2+\rangle}{E_0-E_{l\alpha}}. \tag{47}$$

The energy eigenvalue E_k is related to a root λ by

$$E_k=(\hbar^2/2m)k^2+\lambda. \tag{48}$$

We now examine in detail the matrix elements which occur in the sums L, M, and N above. We note first the selection rules on $\langle r+|\mathbf{p}|l\alpha j\rangle$; $\mathbf{p}$ is a vector and transforms as the representation Γ_{15}^-. The direct product

$$\Gamma_{25}^+\times\Gamma_{15}^-=\Gamma_{12}^-+\Gamma_{15}^-+\Gamma_2^-+\Gamma_{25}^-, \tag{49}$$

so that only the four representations on the right can perturb the valence band edge. The approximate order in energy of the several representations at $\mathbf{k}=0$ in germanium is shown in Fig. 7, based on calculations by F. Herman.[35] It is seen that the states Γ_2^-, Γ_{15}^-, and Γ_{12}^- in the conduction band are likely to provide the most effective perturbations on the state Γ_{25}^+ under consideration.

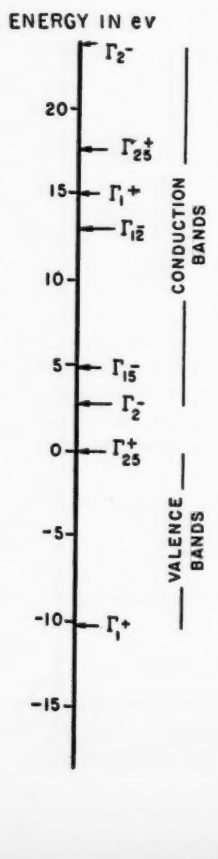


FIG. 7. Proposed order of the energy levels at $\mathbf{k}=0$ in germanium.

There are a number of relations which simplify the matrix elements. Although we do not intend to calculate the matrix elements, we will learn more from the energy surfaces as determined experimentally if we can simplify the expressions for the matrix elements. We first observe that the sum over representations α in L above need be carried out only over the representations Γ_2^- and Γ_{12}^-, as the other matrix elements are seen to vanish on examining the reflection properties of the integrands over the basal planes. For example, we know $\epsilon_1^+\sim yz$ and $p_x\sim x$, so $\epsilon_1^+p_x\sim xyz$. Reference to character tables shows that the characters of the representations Γ_{15}^-, Γ_{25}^- under reflections in the basal planes are positive, while xyz changes sign on reflection; therefore, the corresponding matrix elements vanish. We note the operation JC_4^2 is equivalent to a reflection. In a similar way we see the sum over representations α in M above need be carried out only over representations Γ_{15}^- and Γ_{25}^-.

We now show that L, M, N can be expressed in terms of a single matrix element for each representation, so the sums are reduced essentially to sums over the band index; in practice only one band is expected to contribute significantly. For the Γ_2^- representation, we define

$$F=\frac{\hbar^2}{m^2}\sum_l\frac{|\langle 1+|p_x|\beta_l^-\rangle|^2}{E_0-E_l}, \tag{50}$$

where β_l^- belongs to the one-dimensional representation Γ_2^-. For the Γ_{12}^- representation, we define

$$G=\frac{\hbar^2}{m^2}\sum_l\frac{|\langle 1+|p_x|\gamma_{1l}^-\rangle|^2}{E_0-E_l}. \tag{51}$$

[35] F. Herman (private communication); we are indebted to Dr. Herman for his cooperation on this and other occasions.

We choose γ_1, γ_2 so that the group elements are represented by unitary matrices; this is not done by Von der Lage and Bethe. For example, we take $\gamma_1=x^2+\omega y^2+\omega^2 z^2$ and $\gamma_2=x^2+\omega^2 y^2+\omega z^2$ as a pair of functions transforming according to Γ_{12}^+, where $\omega^3=1$. If we denote $\langle 1+|p_x|\gamma_1^-\rangle$ by R, we have

$$\langle 1+|p_x|\gamma_2^-\rangle=-R, \tag{52}$$

as is seen on rotation by $\pi/2$ about the x-axis. By similar considerations we may show that

$$\langle 2+|p_y|\gamma_1^-\rangle=-\omega^2\langle 2+|p_y|\gamma_2^-\rangle=\omega R. \tag{53}$$

Using (52), we have

$$L=F+2G. \tag{54}$$

For Γ_{15}^- representation, we define

$$H_1=\frac{\hbar^2}{m^2}\sum_l \frac{|\langle 1+|p_y|\delta_{3l}^-\rangle|^2}{E_0-E_l}, \tag{55}$$

where δ_3^- belongs to Γ_{15}^-. The matrix elements with δ_1^-, δ_2^- vanish, as seen by their behavior on reflection in the appropriate basal planes. For the Γ_{25}^- representation, we define

$$H_2=\frac{\hbar^2}{m^2}\sum_l \frac{|\langle 1+|p_y|\epsilon_{3l}^-\rangle|^2}{E_0-E_l}, \tag{56}$$

where ϵ_3^- belongs to Γ_{25}^-; the matrix elements with ϵ_1^-, ϵ_2^- vanish. We have the result

$$M=H_1+H_2. \tag{57}$$

In the sum N all representations appear which satisfy the selection rule (49). The contribution from Γ_2^- is simply F, as

$$\langle 1+|p_x|\beta^-\rangle=\langle 2+|p_y|\beta^-\rangle, \tag{58}$$

by reflection in the (110) plane. The contribution from Γ_{12}^- is $-G$, using Eqs. (52) and (53). The contribution from Γ_{15}^- is H_1, by reflection in the (110) plane. The contribution from Γ_{25}^- is $-H_2$, by reflection. Thus,

$$N=F-G+H_1-H_2. \tag{59}$$

It may be possible to neglect H_2 in silicon and germanium because of remoteness from the valence band.

We must now include the effect of the spin-orbit interaction, which splits the valence band edge into two levels, the upper level being fourfold degenerate ($p_{\frac{3}{2}}$) and the lower level being twofold degenerate ($p_{\frac{1}{2}}$). Our original band had a total degeneracy of $3\times 2=6$, the factor two arising from the two possible orientations of the electron spin. The diamond structure has a center of inversion; it may be shown that each band is doubly degenerate; that is, for a given energy and given k there will be two states. In our original work we incorporated the spin-orbit interaction in the problem using the results of Elliott.[36] We believe now, however, that the results will be more accessible to experimentalists and more closely related to the band energy calculations of Herman and others if presented in terms of a transformation from the $\epsilon_i m_s$ representation to the Jm_J representation.[37]

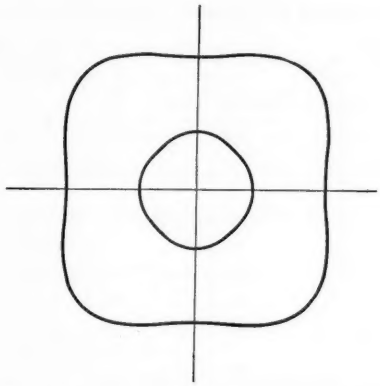


FIG. 8. Figures of constant energy in the (100) plane of k-space for the two fluted energy surfaces which are degenerate at the valence band edge; constants as for germanium.

We note first that to a good approximation we need only be concerned with the transformation of the initial unperturbed states ϵ_i^+ belonging to the representation Γ_{25}^+. A sum of the form

$$\sum_{\alpha j}\langle r+|\mathcal{H}'|\alpha j\rangle\langle\alpha j|\mathcal{H}'|s+\rangle,$$

as in Eq. (46), is invariant under a unitary transformation of the states $|\alpha j\rangle$. If we may neglect the changes in the energy denominators $E_0-E_{l\alpha}$ caused by possible spin-orbit splitting of the states $|\alpha j\rangle$, it follows that the values of the matrix elements $\langle r|\mathcal{H}'|s\rangle$ as in Eq. (43) are not altered by a transformation of the states $|\alpha j\rangle$ to diagonalize the spin-orbit interaction. We may restrict ourselves to spin-orbit effects on the initial states ϵ_i^+.

If we represent the 3×3 matrix in Eq. (46) by Y, the corresponding 6×6 matrix in the $\epsilon_i^+ m_s$ representation is, symbolically,

$$\begin{pmatrix} Y & 0 \\ 0 & Y \end{pmatrix}. \tag{60}$$

We wish to add to the perturbation the spin-orbit interaction,

$$\mathcal{H}_{so}=\frac{\hbar}{4m^2c^2}[\nabla V\times\mathbf{p}]\cdot\boldsymbol{\sigma}, \tag{61}$$

and then to diagonalize the energy matrix with respect to the spin-orbit perturbation; the Jm_J representation is diagonal in the spin-orbit interaction.

The transformed secular equation is, in terms of the

[36] R. J. Elliott, Phys. Rev. 96, 266 (1954).

[37] This type of approach was carried out first by E. N. Adams II (unpublished).

matrix elements H_{ij} defined by Eq. (43),

$$\begin{vmatrix} \frac{H_{11}+H_{22}}{2}-\lambda & -\frac{H_{13}-iH_{23}}{\sqrt{3}} & -\frac{H_{11}-H_{22}-2iH_{12}}{2\sqrt{3}} & 0 & -\frac{H_{13}-iH_{23}}{\sqrt{6}} & -\frac{H_{11}-H_{22}-2iH_{12}}{\sqrt{6}} \\ -\frac{H_{13}+iH_{23}}{\sqrt{3}} & \frac{4H_{33}+H_{11}+H_{22}}{6}-\lambda & 0 & -\frac{H_{11}-H_{22}-2iH_{12}}{2\sqrt{3}} & -\frac{H_{11}+H_{22}-2H_{33}}{3\sqrt{2}} & \frac{H_{13}-iH_{23}}{\sqrt{2}} \\ -\frac{H_{11}-H_{22}+2iH_{12}}{2\sqrt{3}} & 0 & \frac{4H_{33}+H_{11}+H_{22}}{6}-\lambda & \frac{H_{13}-iH_{23}}{\sqrt{3}} & \frac{H_{13}+iH_{23}}{\sqrt{2}} & \frac{H_{11}+H_{22}-2H_{33}}{3\sqrt{2}} \\ 0 & -\frac{H_{11}-H_{22}+2iH_{12}}{2\sqrt{3}} & \frac{H_{13}+iH_{23}}{\sqrt{3}} & \frac{H_{11}+H_{22}}{2}-\lambda & \frac{H_{11}-H_{22}+2iH_{12}}{\sqrt{6}} & -\frac{H_{13}+iH_{23}}{\sqrt{6}} \\ -\frac{H_{13}+iH_{23}}{\sqrt{6}} & -\frac{H_{11}+H_{22}-2H_{33}}{3\sqrt{2}} & \frac{H_{13}-iH_{23}}{\sqrt{2}} & \frac{H_{11}-H_{22}-2iH_{12}}{\sqrt{6}} & \frac{H_{11}+H_{22}+H_{33}}{3}-\lambda-\Delta & 0 \\ -\frac{H_{11}-H_{22}+2iH_{12}}{\sqrt{6}} & \frac{H_{13}+iH_{23}}{\sqrt{2}} & \frac{H_{11}+H_{22}-2H_{33}}{3\sqrt{2}} & -\frac{H_{13}-iH_{23}}{\sqrt{6}} & 0 & \frac{H_{11}+H_{22}+H_{33}}{3}-\lambda-\Delta \end{vmatrix} = 0. \quad (62)$$

In Eq. (62), Δ denotes the spin-orbit splitting of the $p_{\frac{1}{2}}$, $p_{\frac{3}{2}}$ levels. As all H_{ij} are of order k^2, we may approximate the determinant by considering only the elements in the 4×4 block in the upper left corner and in the 2×2 block in the lower right corner. The elements in the two 2×4 strips neglected in this approximation affect the roots only in the order k^4/Δ. The roots of the 4×4 are

$$E(k)=Ak^2\pm[B^2k^4+C^2(k_x^2k_y^2+k_y^2k_z^2+k_z^2k_x^2)]^{\frac{1}{2}}, \quad (63)$$

where

$$\begin{aligned} A&=\tfrac{1}{3}(L+2M)+\hbar^2/2m; \\ B&=\tfrac{1}{3}(L-M); \\ C^2&=\tfrac{1}{3}[N^2-(L-M)^2]. \end{aligned} \quad (64)$$

Each root occurs twice, so that each of the two bands is double; this degeneracy results from the inversion symmetry element of the diamond structure, and is presumably lifted in the zinc blende structure, which includes InSb and other 3–5 semiconductors. The energy surfaces described by Eq. (63) are nonspherical for $C\neq 0$, and are known as fluted or warped surfaces. In Fig. 8 we have plotted in the (100) plane in k-space lines of constant energy for the surfaces in germanium.

The roots of the 2×2 block in Eq. (62) are

$$E(k)=-\Delta+Ak^2, \quad (65)$$

where the constant A is identical with that in Eq. (64) if the spin-orbit splitting Δ may be neglected in comparison with the relevant energy denominators, which are of the order of the forbidden energy gap. This approximation is likely to be satisfactory in silicon, where Δ may be of the order of 0.04 ev, but in germanium Δ is thought to be about 0.3 ev, according to the analysis of Kahn[38] of infrared absorption results in p-Ge. It should be noted that if Δ in silicon is indeed of the order of 0.04 ev, our quadratic expression (63) for the band edge may not be an adequate approximation to describe carriers in thermal equilibrium at room temperature.

[38] A. Kahn, Phys. Rev. **97**, 1647 (1955); thesis, Berkeley, 1954 (unpublished).

The energies near $\mathbf{k}=0$ of other states not split by the spin-orbit interaction are

$$E(\Gamma_1^{\pm})=\frac{\hbar^2k^2}{2m}\left(1+\frac{2}{m}\sum_{\Gamma_{15}^{\mp}}\frac{|\langle\alpha^{\pm}|p_x|\delta_{1l}^{\mp}\rangle|^2}{E_0-E_l}\right); \quad (66)$$

$$E(\Gamma_2^{\pm})=\frac{\hbar^2k^2}{2m}\left(1+\frac{2}{m}\sum_{\Gamma_{25}^{\mp}}\frac{|\langle\beta^{\pm}|p_x|\epsilon_{1l}^{\mp}\rangle|^2}{E_0-E_l}\right); \quad (67)$$

$$E(\Gamma_{12}^{\pm})=k^2(\hbar^2/2m+J+K)\pm(J-K)[k^4-3(k_x^2k_y^2+k_y^2k_z^2+k_z^2k_x^2)]^{\frac{1}{2}}; \quad (68)$$

where

$$\begin{aligned} J&=\frac{\hbar^2}{m^2}\sum_{\Gamma_{15}^{\mp}}\frac{|\langle\gamma_1^{\pm}|p_x|\delta_{1l}^{\mp}\rangle|^2}{E_0-E_l}, \\ K&=\frac{\hbar^2}{m^2}\sum_{\Gamma_{25}^{\mp}}\frac{|\langle\gamma_1^{\pm}|p_x|\epsilon_{1l}^{\mp}\rangle|^2}{E_0-E_l}. \end{aligned} \quad (69)$$

In germanium, where the Γ_2^- state is believed to be the lowest conduction band state, the energy near $\mathbf{k}=0$ is

$$E(\Gamma_2^-)\cong k^2[(\hbar^2/2m)+|F|]. \quad (71)$$

Using the experimental values of the constants [see Eq. (81)], this gives $m^*/m\cong 0.034$. Estimates of the effective masses for other higher conduction states at $k=0$ seem unjustified, as the perturbations on these states will include important terms other than the Γ_{25}^+ valence state.

The fluted or warped quality of the energy surfaces near the valence band edge has a complicated effect on the cyclotron resonance frequency. There is no longer, as with the ellipsoidal surfaces, a single cyclotron frequency for a given orientation of the static magnetic field relative to the axes of the energy surface, but there is now a distribution of resonance frequencies. We give a discussion of the distribution on the assumption that the quantum numbers involved are sufficiently high so that a semiclassical treatment is valid; the quantum theory has been discussed by Luttinger and Kohn,[39] and they find departures from the classical theory at low quantum numbers.

[39] J. M. Luttinger, and W. Kohn Phys. Rev. **97**, 869 (1955).

We suppose that we have a general effective mass Hamiltonian $\mathcal{H}(\mathbf{P})$, where $\mathbf{P}=\hbar\mathbf{k}$. The magnetic field in a classical limit does not change the energy of a particle moving on the energy surface, nor does it change the projection P_H of the $\mathbf{P}$ vector along the magnetic field direction. The motion of the particle is confined to the region in dP_H at P_H and with energy in dE at E; the region has been called a tube by Shockley.[40] The effective mass for cyclotron resonance on any closed tube has been given by Shockley. From Eq. (33) we have $cdP=ev_\alpha Hdt$, where v_α is the scalar magnitude of the projection of the group velocity $\mathbf{v}$ on a plane perpendicular to the magnetic field. Then

$$\oint \frac{cdP}{eHv_\alpha}=\frac{2\pi}{\omega_c}, \tag{72}$$

where ω_c is the fundamental angular frequency of the motion. Higher harmonics are present generally, but are not derived in our present analyses. We define a tube mass m^* such that $\omega_c=eH/m^*c$; thus,

$$m^*=\oint dP/(2\pi v_\alpha). \tag{73}$$

It is always possible, of course, to work directly with the equations of motion, but we have found the integral expression for the mass to be quite convenient. We should emphasize that this equation has been derived in what is essentially a classical limit; the fact that the experimental results appear to be more or less independent of the rf power over a wide range gives us some confidence in the equation; variations in temperature by a factor of two also do not have obvious effects on the positions of the resonance lines.

The cyclotron tubes with $k_H\approx 0$ have the important property that their effective mass remains unchanged as the particle is accelerated and the orbit opens out under the influence of the rf electric field. It is likely, particularly in conditions of high rf power, that the orbits of small k_H have an important effect on the resonance line arising from a fluted energy surface. In other words, the distribution function may be changed by the rf field so as to emphasize small k_H.

We introduce for $\mathbf{k}$ a cylindrical coordinate system, k_H, ρ, ϕ, with k_H parallel to the applied magnetic field; ρ is the radial coordinate in the plane in k-space perpendicular to k_H. By an elementary transformation of Eq. (73), we have

$$m^*=\frac{\hbar^2}{2\pi}\oint \frac{\rho d\phi}{(\partial E/\partial\rho)}. \tag{74}$$

The application of this result to the valence band edge of silicon and germanium is generally formidable. The result is fairly tractable for H parallel to a (110) plane and for the equatorial tubes $k_H=0$. We have, under these restrictions,

$$m^*=(\hbar^2/\pi)\int_0^{\pi/2}\frac{d\phi}{A\pm\{B^2+\frac{1}{4}C^2[1+g(\phi)]\}^{\frac{1}{2}}}, \tag{75}$$

where A, B, C are defined by Eq. (63), and

$$g(\phi)=-(3\cos^2\theta-1)[(\cos^2\theta-3)\cos^4\phi+2\cos^2\phi], \tag{76}$$

with θ the angle the magnetic field makes with the [100] direction.

An expansion in power of $g(\phi)$ gives

$$m^*=\frac{\hbar^2}{2}\frac{1}{A\pm[B^2+(C/2)^2]^{\frac{1}{2}}} \times\left\{1\pm\frac{C^2(1-3\cos^2\theta)^2}{64[B^2+(C/2)^2]^{\frac{1}{2}}\{A\pm[B^2+(C/2)^2]^{\frac{1}{2}}\}}+\cdots\right\}. \tag{77}$$

This result is exact for the [111] direction; in other directions the contribution of the next term in the expansion is not greater than about 1 percent in silicon and germanium.

We have evaluated the constants A, B, C by making a fit of m^* as given by Eq. (77) to the experimental data on the two cyclotron resonance lines associated with the valence band. This procedure is justified if the position of the center of the resonance line is given approximately by the carriers which have $k_H\approx 0$. There are two arguments which support this procedure; the first argument given above is that the orbits near $k_H=0$ maintain their frequency constant as they are accelerated outward; the second argument is that m^* is fairly independent of k_H, except for high k_H, which are discriminated against by a geometrical factor in the density of states. In Fig. 9 we give the results of calculations of m^* *vs* k_H for germanium in the [100] and [111] directions. The contributions of high k_H are principally in one wing of the resonance line.

In Fig. 10 we give a plot of the experimental points for holes in germanium at 4°K as a function of the angle between the direction of the static magnetic field in a (110) plane and a [001] direction lying in the plane. The constants A, B, C in the expression (63),

$$E(k)=Ak^2\pm[B^2k^4+C^2(k_x^2k_y^2+k_y^2k_z^2+k_z^2k_x^2)]^{\frac{1}{2}}, \tag{78}$$

are determined from the experimental data, using Eq. (77). The theoretical curve in the figure is calculated using the values

$$\begin{aligned} A&=-(13.0\pm 0.2)(\hbar^2/2m);\\ |B|&=(8.9\pm 0.1)(\hbar^2/2m);\\ |C|&=(10.3\pm 0.2)(\hbar^2/2m). \end{aligned} \tag{79}$$

These values represent our best fit for germanium. Dexter, Zeiger, and Lax[41] have reported $A=-13.6(\hbar^2/$

[40] W. Shockley, Phys. Rev. **79**, 191 (1950).

[41] Dexter, Zeiger, and Lax, Phys. Rev. **95**, 557 (1954).

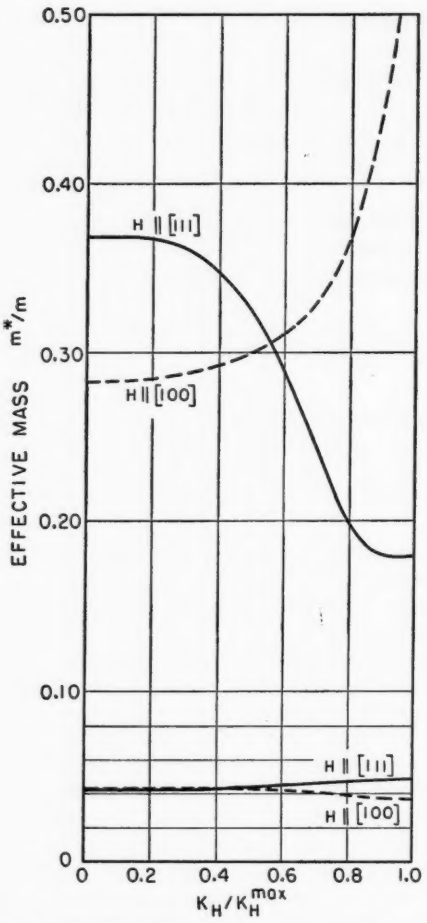


FIG. 9. Calculated values of m^* *vs* the component k_H of the k-vector along the direction of the static magnetic field, for the [100] and [111] directions in germanium.

$2m)$; $|B|=9.1(\hbar^2/2m)$; $|C|=11.2(\hbar^2/2m)$. Our error estimates in (79) represent the scatter of the experimental points about the theoretical curve; because of the use of the assumption $k_H\approx 0$, the correct constants may possibly lie outside of the indicated limits. An attempt by us to take into account the distribution of k_H gives the following approximate constants: $A=-(13.2\pm 0.1)(\hbar^2/2m)$; $|B|=(8.9\pm 0.05)(\hbar^2/2m)$; $|C|=(10.6\pm 0.2)(\hbar^2/2m)$.

In Fig. 11 we give a plot of the experimental points for holes in silicon at 4°K. The theoretical curve is calculated using the values

$$\begin{aligned} A&=-(4.1\pm 0.2)(\hbar^2/2m);\\ |B|&=(1.6\pm 0.2)(\hbar^2/2m);\\ |C|&=(3.3\pm 0.5)(\hbar^2/2m). \end{aligned} \tag{80}$$

Dexter and Lax[42] have reported $A=-4.0(\hbar^2/2m)$; $|B|=1.3(\hbar^2/2m)$; $|C|=3.6(\hbar^2/2m)$. An attempt by us to take into account the distribution of k_H gives the following approximate constants: $A=-(4.0\pm 0.2)(\hbar^2/2m)$; $|B|=(1.1\pm 0.5)(\hbar^2/2m)$; $|C|=(4.0\pm 0.5)(\hbar^2/2m)$.

Using relations (54), (57), (59), (64), together with the experimental results from (79) and (80), the values of the sums over matrix elements in germanium are

$$\begin{aligned} L&=-31.8(\hbar^2/2m); & F&=-28.6(\hbar^2/2m);\\ M&=-5.1(\hbar^2/2m); & G&=-1.6(\hbar^2/2m);\\ N&=-32.1(\hbar^2/2m); & H_1&=-5.1(\hbar^2/2m);\\ & & H_2&=0; \end{aligned} \tag{81}$$

and, in silicon,

$$\begin{aligned} L&=-1.9(\hbar^2/2m); & F&=-1.2(\hbar^2/2m);\\ M&=-6.7(\hbar^2/2m); & G&=-0.4(\hbar^2/2m);\\ N&=-7.5(\hbar^2/2m); & H_1&=-6.7(\hbar^2/2m);\\ & & H_2&=0. \end{aligned} \tag{82}$$

If we neglect the spin-orbit splitting of the valence band edge in comparison with the other energy denominators, the effective mass of the $p_{\frac{1}{2}}$-band should be, from Eq. (65), $m^*\cong 0.08m$ in germanium and $m^*=0.25m$ in silicon. The value for germanium is in satisfactory agreement with Kahn's interpretation of the infrared absorption spectrum of p-Ge. These constants are not uniquely determined from A, B, and C, because Eqs. (64) are not linear. The constant H_2 was assumed zero due to the presumed remoteness of the Γ_{25}^- state from the valence band. The choice of constants above was made because the sums over matrix elements are all negative, as expected if the conduction

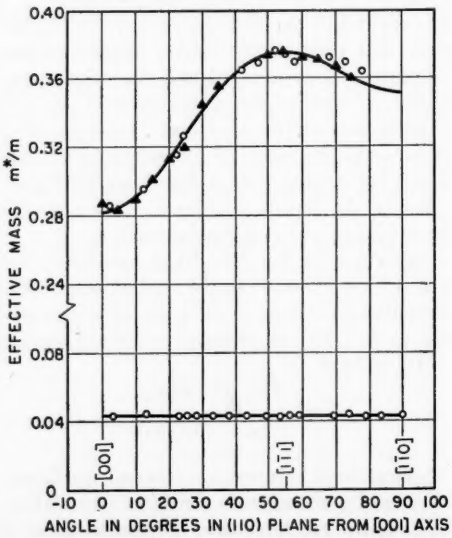


FIG. 10. Effective mass of holes in germanium at 4°K for magnetic field directions in a (110) plane; the theoretical curves are obtained from Eq. (77), using the constants in Eq. (79).

[42] R. N. Dexter and B. Lax, Phys. Rev. 96, 223 (1954).

states furnish the main perturbation. In silicon there is another set of solutions in which all sums over matrix elements are negative, but the order of magnitude of the constants seemed unlikely. The constants above are in line with the model proposed by Herman.[43] We note that perturbations with Γ_2^- are dominant in germanium, and with Γ_{15}^- are dominant in silicon.

We have analyzed, using Eq. (8), the line widths observed at 4°K in the specimens which gave the sharpest lines. The relaxation times for electrons were approximately isotropic, with $\tau(\text{Ge}) \cong 6\times10^{-11}$ sec and $\tau(\text{Si}) \cong 7\times10^{-11}$ sec. The effective relaxation times for the light mass hole resonances were $\tau(\text{Ge}) \cong 7\times10^{-11}$ and $\tau(\text{Si}) \cong 7\times10^{-11}$ sec; for the heavy mass hole resonances, $\tau(\text{Ge}) \gtrsim 5\times10^{-11}$ sec and $\tau(\text{Si}) \gtrsim 6\times10^{-11}$ sec. These data were taken with optical excitation of carriers. The lines did not appear to sharpen appreciably on pumping to 2°K, but the specimens may possibly have been at a higher temperature. At rf power levels below the ionization limit it did not appear that the widths were dependent on the rf power levels in the range covered. The line shapes of the electron resonances appeared to be approximately Gaussian, while the elementary theory predicts a Lorentzian shape.

Several remarks can be made about the relative intensity of the light and heavy mass hole resonance lines. At the resonance maximum the energy losses are proportional, according to Eq. (15), to the static conductivity $\sigma_0 = Ne^2\tau/m^*$. Therefore,

$$\frac{I(l)}{I(h)} = \frac{N_l}{N_h}\frac{m_h}{m_l}\frac{\tau_l}{\tau_h}. \tag{83}$$

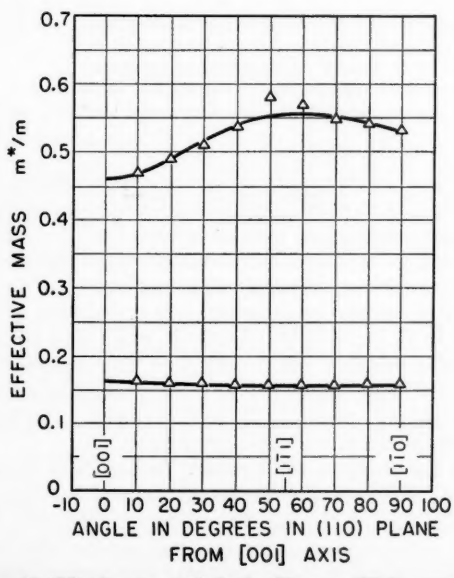


FIG. 11. Effective mass of holes in silicon at 4°K for magnetic field directions in a (110) plane; the theoretical curves are obtained from Eq. (77), using the constants in Eq. (80).

[43] F. Herman, Phys. Rev. **95**, 847 (1954).

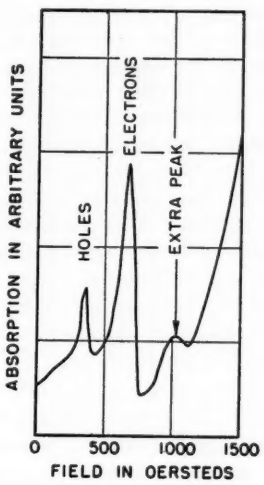


FIG. 12. Possible extra line, as indicated in germanium at 55° from a [001] direction in a (110) plane.

If the population ratio N_l/N_h of the two bands is determined by considerations of thermal equilibrium, we have

$$N_l/N_h = (m_l/m_h)^{\frac{3}{2}}, \tag{84}$$

as the volume in phase space corresponding to an energy range ΔE is proportional to $(m^*)^{\frac{3}{2}}$. Thus,

$$I(l)/I(h) = (m_l/m_h)^{\frac{3}{2}}(\tau_l/\tau_h). \tag{85}$$

In silicon and germanium it appears that $\tau_l \cong \tau_h$, so we would expect, roughly,

$$\text{Ge: } I(l)/I(h) \cong \tfrac{1}{8}; \quad \text{Si: } I(l)/I(h) \cong \tfrac{1}{2}; \tag{86}$$

as the average mass ratio is about $\frac{1}{8}$ for germanium and $\frac{1}{3}$ for silicon.

The integrated intensity will be proportional to the product of the peak intensity and the line width $\Delta\omega_c = 1/\tau = (e/m^*c)\Delta H$, so that the integrated intensity ratio is

$$g(l)/g(h) = (m_l/m_h)^{\frac{3}{2}}. \tag{87}$$

We therefore expect

$$\text{Ge: } g(l)/g(h) \cong 1/20; \quad \text{Si: } g(l)/g(h) \cong 1/5. \tag{88}$$

The experimental ratios are consistent with these estimates.

6. FURTHER REMARKS

Kip[44] at the Amsterdam Conference reported the tentative observation in germanium of several extra lines, that is, lines which cannot be assigned to the hole or electron band edge energy surfaces. The observational situation on the extra lines is rather unsatisfactory at the moment, as one of the lines (the one shown in Fig. 12) is feeble and does not often appear. The second extra line (not shown) is erratic in appearance, and when it does appear it is rather too strong to be creditable; we are inclined to believe that at least in our own work the appearance of the second

[44] A. F. Kip, Physica **20**, 813 (1954).

line may often be accounted for by a slight misorientation of the specimen, as a 5° misorientation can remove the degeneracy of one of the electron resonance lines in the (110) plane by a splitting of as much as 200 to 300 oersteds for some directions. There are several mechanisms which one might invoke for the production of extra lines, including (a) the possibility of resonance on excited bands or near higher local minima on the usual band; (b) partial breakdown of the selection rule $\Delta n=\pm 1$ on the fluted surfaces, as suggested privately by Dexter, Lax, and Zeiger; (c) nonclassical effects at low temperatures, as hinted at by Kohn and Luttinger[45]; (d) distortion of the form of the energy surfaces in the valence band at small k as a result of the Zeeman splitting of the band edge states; and (e) if the plot of m^* *vs* k_H should be horizontal at several separated k_H values, extra lines should appear.

We have recently observed cyclotron resonance of electrons and holes in InSb, and we have a preliminary indication of cyclotron resonance in InAs. Details of this work will be published separately.

[45] W. Kohn and J. M. Luttinger, Phys. Rev. **96**, 529 (1954).

ACKNOWLEDGMENTS

We have had a great deal of valuable assistance from many persons and organizations in this work. Financial support is gratefully acknowledged from the U. S. Office of Naval Research, the U. S. Signal Corps, the National Science Foundation, and the Pittsburgh Plate Glass Foundation. We are especially indebted for the supply of semiconductor crystals to the Bell Telephone Laboratories, to Sylvania Electric Products, Inc., and to the Westinghouse Research Laboratories. Liquid helium was kindly furnished by Professor W. F. Giauque and Dr. D. N. Lyons. The crystals were oriented by Prof. J. Washburn. J. Ubbink, G. Feher, and Glen Wagoner assisted with the measurements. We wish to express our thanks to the Lincoln Laboratory group associated with Dr. B. Lax for their friendly exchange of information on their work. We have profited from correspondence and conversations with E. N. Adams II, P. Aigrain, W. B. Brattain, E. Burstein, E. Conwell, R. E. Davis, R. J. Elliott, H. Y. Fan, R. Fletcher, M. J. E. Golay, F. Herman, C. Herring, R. Longini, F. J. Morin, H. M. O'Bryan, F. Seitz, W. Shockley, and A. F. Siefert.

PHYSICAL REVIEW VOLUME 98, NUMBER 2 APRIL 15, 1955

Measurement of Shot Noise in CdS Crystals

C. I. SHULMAN
RCA Laboratories, Princeton, New Jersey
(Received September 21, 1954)

The noise power spectrum associated with photoconduction current in CdS crystals with indium electrodes is found to flatten off at low frequencies at a value that corresponds closely to the noise inherent in the photon absorption process itself plus that associated with the random nature of the carrier recombination process. It is found that the noise power is not a unique function of the photoconduction current, but varies as the square of the applied voltage, and linearly with light intensity, as suggested by a simple model not unlike that of the photomultiplier.

MEASUREMENTS of noise associated with the passage of current through semiconducting materials have led to a variety of theories,[1–4] most of which center about boundary layer phenomena. The essential observations that these models seek to derive are (1), the large excess of noise relative to thermal and (2) a $1/f$ spectrum down to extremely low frequencies. Most of the models are complicated and have the flexibility to account for a wide range of spectra and often do, in fact, give a fairly good fit with observation. There are simple models[5,6] involving processes within the body of the semiconductor itself which yield spectra that flatten off at low frequencies and give noise levels much lower than those reported by most observers. These models are rather straightforward and simple, and represent a possible reference for the study of inner processes in semiconductors, for if the strong fluctuations associated with boundary layer phenomena could be eliminated, one would gain a useful tool for the study of current flow in solids. It is the purpose of this note to describe work wherein it was found possible to make noiseless ohmic contact to CdS crystals such that the noise associated with the passage of current through the crystal could be interpreted in terms of processes within the CdS crystal itself.[7]

In the study of the electrical properties of crystals there is always the important general question of separating out the effects of contact electrodes and their interaction with the material under study. Such interactions often cause nonlinear volt-ampere characteristics and nonlinear potential distributions within the body of crystals. The use of gallium[8] or indium as

[1] W. Schottky, Phys. Rev. **28**, 74 (1926).
[2] G. G. McFarlane, Proc. Phys. Soc. (London) **59**, 366 (1947).
[3] A. Van D. Ziel, Physica **16**, 359 (1950).
[4] W. M. Buttler, Ann. Physik. **11**, 362 (1953).
[5] B. Davydov and B. Gurevich, J. Phys. (U.S.S.R.) **7**, 138 (1943).
[6] J. H. Gisolf, Physica **15**, 825 (1949).
[7] Shulman, Smith, and Rose, Phys. Rev. **92**, 857(A) (1953).
[8] R. W. Smith, Phys. Rev. **92**, 857 (1953).

electrodes for CdS crystal studies changed the picture considerably, for not only was it found possible to obtain linear volt-ampere characteristics down to the smallest measurable voltages, but linear potential distributions within the body of the crystal were also observed. It was further found that noise arising from current flow through such crystals was considerably less than that observed in crystals with gold or silver contacts of various types. In fact, some CdS crystals showed noise that was closely attributable to the random absorption of photons within the body of the crystal.

To develop the simplest of models for fluctuations in photoconduction, we begin with a very simple description of the photoconduction process which involves one mobile carrier, no traps, and need not refer to a model for the material at all. Quantum gain, time constant, frequency response, and noise spectrum are all interrelated in a simple way, independent of crystal model, as follows. From the assumption of unity quantum efficiency in the excitation process, and space charge neutrality in the materials, one can immediately write the quantum gain[9] as

$$G_0=\tau/T, \tag{1}$$

where G_0 is the ratio of the number of electrons passing through the crystal to the number of photons absorbed by the crystal, τ is the carrier lifetime, and T is the transit time through the crystal. The frequency response of such a model is given by

$$1/(1+\omega^2\tau^2)^{\frac{1}{2}}, \tag{2}$$

while the mean square noise current per unit band width is given by[10]

$$\langle i_n^2\rangle_{Av}=4eI_0G_0/(1+\omega^2\tau^2) \tag{3}$$

where I_0 is the crystal photocurrent.

This noise spectrum can be derived, except for a factor of 2, in a simple way by thinking of the random absorption of photons as being a source of shot noise with a uniform spectrum which is modified by the crystal response given by Eq. (2). We would then have

$$\langle i_n^2\rangle_{Av}=2eFeG_0^2/(1+\omega^2\tau^2), \tag{4}$$

where F=average rate of absorption of photons. But $FeG_0=I_0$, so (4) reduces to (3), with a factor 2 instead of 4. Since the spectrum associated with fluctuations in gain is in this case identical with that due to random arrival of photons, gain fluctuation can be taken into account by the simple factor of 2.

[9] A. Rose, R.C.A. Rev. 12, 362 (1951).

[10] This expression for noise current is derived by considering random passage of square pulses of current of height e/T, and duration S with probability distribution $(1/\tau)e^{-S/\tau}$. The usual derivations give Eq. (3) with a factor 2 instead of 4. The addition factor 2 comes from fluctuations in S as pointed out by D. O. North in a private communication. Davydov and Gurevich obtained Eq. (3) with the factor 4 for small lifetimes, and would have obtained the same result for all lifetimes if they had taken into account the correlation between carriers entering at the cathode and those leaving at the anode.

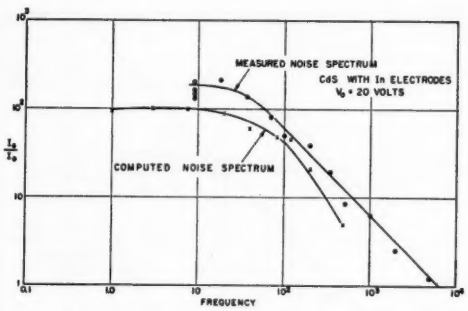


FIG. 1. Comparison of observed noise spectrum with noise spectrum computed from photon shot noise, crystal gain, and frequency response.

It is most convenient to refer this noise spectrum to that of an ordinary temperature limited diode in which

$$\langle i_n^2\rangle_{Av}=2eI_e(f),$$

where I_e is the current in the temperature limited diode to give the same noise as is observed in the crystal. Hence we can write

$$I_e(f)=2G_0I_0/(1+\omega^2\tau^2).$$

For a uniform potential distribution in the crystal

$$G_0=\tau/T=\mu V_0\tau/L^2,$$

where μ=carrier mobility, L=crystal length, and V_0=voltage across the crystal length. Then

$$I_e(f)=2\frac{\mu\tau}{L^2}V_0I_0\frac{1}{1+\omega^2\tau^2}. \tag{5}$$

The essential features of this spectrum are (1) that it flattens off at low frequencies to a limiting value given by $2G_0I_0$, (2) its shape corresponds to the square of the crystal response characteristic, (3) for a given light intensity, $I_e(f)$ is proportional to V_0^2 and (4) for a given V_0, $I_e(f)$ is proportional to I_0. The last two follow from the observed linear dependence of I_0 on V_0, and of conductivity on light intensity. Experimental establishment of the relationships between these observables represents a good foundation from which photoconductors can be studied and on which to build models for the cases where departures from these relations are observed.

Measurements on CdS crystals with indium pellet electrodes show some of the features of the simple spectrum described above. Figure 1 shows a comparison between measured noise spectrum and the noise spectrum computed from photon shot noise, measured crystal gain, and measured frequency response of the crystal. Although there is scatter at the low end of the spectrum, it does flatten off at a value approximately twice that predicted by the simplest of theories. If one were to insert effect due to surface reflection the computed curve would move up toward the measured curve by 20 to 30 percent. This value should be compared with Butler's[4] measurements of the noise inte-

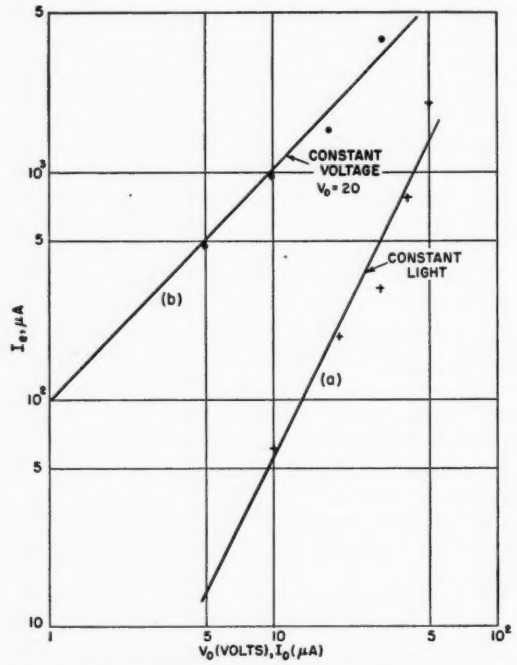


FIG. 2. Dependence of I_e on V_0 and I_0, measured at 40 cps.

grated over the entire spectrum in which he obtains noise powers six to eight powers of ten greater than that corresponding to the integrated noise derived from the simple theory. This suggests that he was dealing with complicated processes probably associated with electrode contacts. The dashed line which is the spectrum computed from measured gain and frequency response, represents the minimum possible noise power for it is due to the randomness of the photon absorption process and fluctuations in gain. The degree to which the measured noise approaches the minimum possible value represents the degree to which all other processes are eliminated. Considering uncertainties in measurements of crystal gain because of uncertainties in surface reflectivities, the closeness with which the leveling off noise power approaches minimum represents a strong argument for the simplest of models for there is not much room for complicated internal processes to add to the noise already inherent in the photon process.

There is however only a weak correspondence between the shapes of the measured noise spectrum and the spectrum derived from the simple model, for the half-power points disagree by a factor two, and the high-frequency end shows $1/f$ law for the noise spectrum and $1/f^2$ for the response characteristic. Although this difference is not understood, it is probably due to trapping processes and close study here may indeed be a fruitful source of information.

Figure 2 shows the dependence of noise power on V_0 and I_0, measured at 40 cps. Square-law dependence on V_0 and linear dependence on I_0 is indicated and

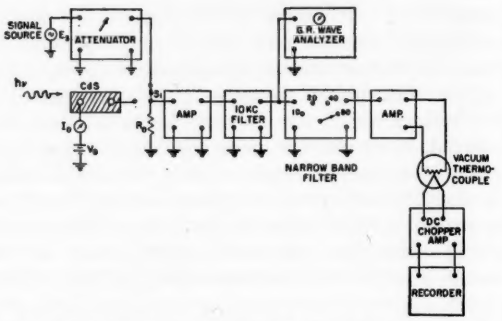


FIG. 3. Schematic arrangement for noise measurements.

corresponds to the simple picture. Curve (a) shows the variation in I_e with V_0 for constant illumination intensity, while curve (b) shows the variation in I_e with I_0, for constant V_0. I_0 is varied by varying the light intensity. Although other experiments give square-law dependence on voltage, the combination of square-law dependence on voltage and linear dependence on current is new,[11] and represents support for the simple model.

The general picture is that although these observations do not rigorously establish this simplest of models, it is believed to have sufficient merit to be useful as a guide in one's thinking on fluctuation processes in photoconductors. It should be pointed out however that not all crystals have shown this behavior. For instance CdS crystals with evaporated indium electrodes which were macroscopically ohmic showed large variations in noise power and in general gave more noise than those with indium pellet electrodes. There were also crystals with indium pellet electrodes which gave high noise.

The measurements were made with the arrangement shown in Fig. 3. The signal source E_s and the attenuator were adjusted to give the same reading on the output meter. The General Radio wave analyzer was used for readings from 100 cps on up. It has a noise band width of about 5 cps over its entire range. At low frequencies the small drift inherent in beat frequency instruments made it necessary to build narrow band filters for making the noise measurement. Twin T feed-back filters were built for 10, 20, 40, and 80 cps with noise bandwidths of 1 cps or less. To integrate out the slow fluctuation, a recorder was used to average each reading over several minutes. The crystal response was taken with a sinusoidally modulated kinescope light source, while the dc gain was measured with a calibrated photocell.

[11] A. Slocum and J. N. Shive [J. Appl. Phys. **25**, 406 (1954)] report measurements in which photocurrent noise in p-n junctions can be attributed to random arrival of photons and is proportional to photocurrent. However, their measurements correspond to unity gain and voltage saturated photocurrents, in contradistinction to the measurements described herein which involve high gains and photocurrents proportional to the applied voltage. Under this condition it is possible to observe two noise powers for a given photocurrent, depending on whether the condition is that of low light and high voltage, or high light and low voltage.

PHYSICAL REVIEW VOLUME 98, NUMBER 2 APRIL 15, 1955

Interaction between Spin Waves and Conduction Electrons in Ferromagnetic Metals*

ELIHU ABRAHAMS
Department of Physics, University of Illinois, Urbana, Illinois
(Received September 23, 1954)

With a view toward understanding ferromagnetic resonance line widths in metals, the relaxation time due to the spin wave-conduction electron interaction has been calculated. The effect due to the interaction of ferromagnetic spins with conduction electron currents is much more important than that due to the interaction of ferromagnetic spins with conduction electron spins. In nickel, for temperatures between 4°K and 300°K, the former process yields a relaxation time $10^{-8}-10^{-7}$ sec, while the latter yields $10^{-3}-10^{-4}$ sec. The interactions do not appear to be strong enough to account for the observed line widths.

I. INTRODUCTION

THE role of relaxation processes in determining the line width in ferromagnetic spin resonance absorption[1] has recently been discussed by several authors.[2–4] According to the discussion of KA,[4] for temperatures less than about one-half the Curie temperature, the line width is determined by spin-spin interactions of magnetic dipolar and pseudodipolar origin. On this view, a calculation of the spin-lattice relaxation involves the assumption that the ferromagnetic spin system is in thermal equilibrium at a definable spin temperature which is higher than the lattice temperature. This is an attractive idea and it greatly simplifies the calculation of spin-lattice relaxation times; furthermore, the experimental evidence lends some degree of support. However, this interpretation leaves the line width still unexplained, for it has not yet been demonstrated quantitatively that spin-spin interactions can lead to the large temperature independent line widths that have been observed at low temperatures.[5] Whether or not spin-spin interactions can be responsible for the line widths at low temperatures must therefore remain an open question. In the calculation which follows, we ascertain that the interaction of the magnetization (ferromagnetic spin system) with the conduction electrons cannot account for the major part of the line width in ferromagnetic metals at low temperatures. The line width must then be due to another mechanism which, according to KA, is the spin-spin interaction.

The direct spin-lattice relaxation time has previously been calculated using a macroscopic interaction between spin-wave and phonon fields.[6] In the present work, a calculation is made, by similar techniques, of the relaxation processes which transfer energy from the ferromagnetic spin system to the lattice by way of the conduction electrons (s-electrons). Since the conduction electron-lattice interaction is so strong (with a characteristic time of the order of 10^{-13} sec),[7] we may consider the s-electrons to be part of the lattice and we then treat the energy transfer as a "spin-electron relaxation." The conduction electrons will be treated as a degenerate electron gas by means of the single particle model in the plane wave approximation. The spin waves will be treated by methods discussed by KA.

It has been shown[8] that in a ferromagnetic conductor, the resonance frequency for the absorption of microwave power is shifted by an amount inversely proportional to the square of the eddy current skin depth δ. In spin-wave language, this means that the only spin waves excited by the rf field have wave vectors κ whose magnitudes lie close to the value $1/\delta$. This is to be compared with the situation in insulators where the microwave excitation is limited to those spin waves having wave numbers κ near zero. In the case of the direct spin-lattice interaction, the difference is insignificant as the phonons of interest have wave numbers very much larger than $1/\delta$, and the previous calculation[9] for the direct interaction which scatters $\kappa=0$ spin waves remains essentially unchanged in the case of metals. In order to determine whether the interaction of spin waves with conduction electrons can account for the line width, it is necessary to evaluate the time constant (relaxation time) of the decay of an excess number of spin waves of wave number $1/\delta$. If this interaction is to account for the width, the relaxation time must be as short as 10^{-9}–10^{-10} sec.

Previous work on spin-wave-electron interactions has been reported by Samoilovich and Yokovleff[10] in connection with electrical resistivity. Kondoh[11] has published a report on a calculation using the dipolar interaction between s- and d-electrons in the spin-wave

* Research supported in part by the Office of Naval Research.
[1] Ferromagnetic resonance is reviewed by C. Kittel, J. phys. radium **12**, 291 (1951); J. H. Van Vleck, Physica **17**, 234 (1951).
[2] P. W. Anderson, Phys. Rev. **88**, 1213 (1952).
[3] N. Bloembergen and S. Wang, Phys. Rev. **93**, 72 (1954).
[4] C. Kittel and E. Abrahams, Revs. Modern Phys. **25**, 233 (1953); henceforth denoted by KA.
[5] N. Bloembergen, Phys. Rev. **78**, 572 (1950). A. F. Kip (private communication).
[6] E. Abrahams and C. Kittel, Phys. Rev. **88**, 1200 (1952). See also reference 3.
[7] It should be mentioned that the relaxation of s-electron spins to the lattice is also expected to be very fast due to the large spin-orbit coupling. See R. J. Elliott, Phys. Rev. **96**, 266 (1954).
[8] C. Kittel and C. Herring, Phys. Rev. **77**, 725 (1950).
[9] J. M. Luttinger and C. Kittel (unpublished). The present author has improved the calculation of Luttinger and Kittel and finds that the spin-lattice relaxation time for $\kappa=0$ spin waves in nickel is of the order of 10^{-5} sec at room temperature.
[10] A. G. Samoilovich and V. A. Yokovleff, Zhur. Eksptl. i. Teoret. Fiz. **22**, 350 (1952).
[11] A. Kondoh, Progr. Theoret. Phys. Japan **10**, 117 (1953).

approximation of Holstein and Primakoff.[12] There is essential agreement between his results and those of a similar calculation which forms Sec. IV of this paper.

II. INTERACTIONS

The interaction operator which we adopt is

$$H_{\text{int}}=H_1+H_2, \tag{1}$$

$$H_1=\frac{e}{mc}\int\frac{\mathbf{M}(\mathbf{r})\cdot(\mathbf{r}-\mathbf{r}_e)\times\mathbf{p}}{|\mathbf{r}-\mathbf{r}_e|^3}e^{-q|\mathbf{r}-\mathbf{r}_e|}d\mathbf{r}, \tag{2}$$

$$H_2=\beta\boldsymbol{\sigma}\cdot\mathbf{H}_e. \tag{3}$$

where $\mathbf{M}(\mathbf{r})$ is the magnetization at the point $\mathbf{r}$, $\mathbf{r}_e$ is the position of a conduction electron, $\mathbf{p}$ its momentum and $\boldsymbol{\sigma}$ its spin. $\mathbf{H}_e$ is the magnetic field due to the ferromagnetic spins and has been given by Herring and Kittel.[13] In Eq. (3), β is the Bohr magneton. H_1 represents an interaction between the ferromagnetic spins and the currents due to the motion of the s-electrons. Furthermore $1/q$ is the screening radius[14] for the magnetic field due to the s-electron currents and is given by $(mc^2/4\pi n_0e^2)^{-\frac{1}{2}}$ where n_0 is the s-electron density. This screening arises from the collective influence of the set of s-electrons and is discussed fully in reference 14. H_2 is the ferromagnetic spin—s-electron spin interaction in the Herring-Kittel approximation. The two interactions H_1 and H_2 will be treated separately since they do not interfere.

The system of conduction electrons is treated as a quantized free electron field with the following convention for the field operators:

$$\begin{aligned}&\text{spin up:}\quad \varphi=b_Ke^{i\mathbf{K}\cdot\mathbf{r}},\\ &\text{spin down:}\quad \psi=d_Ke^{i\mathbf{K}\cdot\mathbf{r}},\end{aligned} \tag{4}$$

where $\mathbf{K}$ is the electron wave vector. Following KA, we have for the spin-wave field:

$$\begin{aligned}M^+(\mathbf{r})&=M_x(\mathbf{r})+iM_y(\mathbf{r})=(2g\beta M_s)^{\frac{1}{2}}\textstyle\sum_\kappa\alpha_\kappa e^{i\boldsymbol{\kappa}\cdot\mathbf{r}},\\ M^-(\mathbf{r})&=M_x(\mathbf{r})-iM_y(\mathbf{r})=(2g\beta M_s)^{\frac{1}{2}}\textstyle\sum_\kappa\alpha_\kappa^*e^{-i\boldsymbol{\kappa}\cdot\mathbf{r}},\end{aligned} \tag{5}$$

where α_κ, b_K, d_K, and their conjugates are annihilation and creation operators and M_s is the saturation magnetization. We work throughout in Gaussian units with a sample of unit volume.

III. INTERACTION WITH S-ELECTRON CURRENTS

The interaction of spin waves with s-electron currents does not cause s-electron spin transitions so that we may consider interactions with s-electrons of one spin only. From Eq. (2), we get the perturbation operator $\mathcal{H}_1=\int\varphi^*H_1\varphi d\mathbf{r}_e$:

$$\mathcal{H}_1=(e\hbar/mc)\textstyle\sum_{KK'}b_K^*b_{K'}\int d\mathbf{r}e^{i(\mathbf{K}'-\mathbf{K})\cdot\mathbf{r}}$$

$$\times\mathbf{M}(\mathbf{r})\cdot\int(\mathbf{R}\times\mathbf{K}'/R^3)e^{i(\mathbf{K}'-\mathbf{K})\cdot\mathbf{R}-qR}d\mathbf{R}. \tag{6}$$

Here we have considered electrons with spin up. The evaluation of the integrals which appear here is facilitated by the expansion of $e^{i(\mathbf{K}'-\mathbf{K})\cdot\mathbf{R}}$ in spherical harmonics. The result is

$$\mathcal{H}_1=(2\pi ie\hbar/mc)(2g\beta M_s)^{\frac{1}{2}}[1-(q/\kappa)\tan^{-1}(\kappa/q)]$$

$$\times\textstyle\sum_{K\kappa}b_K^*b_{K'\kappa}\kappa^{-2}[(\mathbf{K}\times\boldsymbol{\kappa})^+\alpha_{-\kappa}^*+(\mathbf{K}\times\boldsymbol{\kappa})^-\alpha_\kappa],$$

$$\boldsymbol{\kappa}=\mathbf{K}-\mathbf{K}'\neq0, \tag{7}$$

plus another term which does not contribute to the relaxation. The first term in square brackets creates a spin wave of wave vector $-\boldsymbol{\kappa}$ while the second destroys one of wave vector $\boldsymbol{\kappa}$. In order to evaluate the decay rate of spin waves of given wave number, we fix κ and consider the rate of change of the population in the state κ if there is a nonequilibrium excess of spin waves in that state. The nonequilibrium excess is due to the excitation by the microwave field.

The net number of collisions per unit time which transfer energy from the spin system to the s-electrons is given by the kinetic equation

$$\dot{N}_{\text{coll}}=(2\pi/\hbar)\textstyle\sum_{K\kappa}(|H_1^e|^2-|H_1^a|^2)\delta(\epsilon_\kappa+\epsilon_{K'}-\epsilon_K), \tag{8}$$

where H_1^e and H_1^a are those parts of Eq. (7) which emit and absorb, respectively, a spin wave of wave vector $\boldsymbol{\kappa}$, and ϵ_κ, ϵ_K are the energies of a spin wave and an electron of wave vectors $\boldsymbol{\kappa}$, $\mathbf{K}$. Here ϵ_κ is equal to $(2g\beta A/M_s)\kappa^2-g\beta H$ where A is the macroscopic exchange constant and H is the applied constant field.

To set up the relaxation time, we fix κ and identify $\dot{N}_{\text{coll}}$ with $\dot{n}_\kappa$, the rate of change of population in the state κ. Further, we take the matrix elements of the creation and destruction operators and average over the directions of the vector $\mathbf{K}\times\boldsymbol{\kappa}$. The result for $\dot{n}_\kappa$ is

$$\dot{n}_\kappa=(32\pi^3/3)(e^2\hbar g\beta M_s/m^2c^2)$$

$$\times[1-(q/\kappa)\tan^{-1}(\kappa/q)]^2\kappa^{-4}\textstyle\sum_K(\mathbf{K}\times\boldsymbol{\kappa})^2$$

$$\times[(n_\kappa+1)(1-f_{K'})f_K-n_\kappa(1-f_K)f_{K'}]$$

$$\times\delta(\epsilon_\kappa+\epsilon_{K'}-\epsilon_K), \tag{9}$$

$$\boldsymbol{\kappa}=\mathbf{K}-\mathbf{K}'\neq0.$$

Here n_κ is the occupation number of the spin-wave state κ and f_K is the probability of finding an s-electron in the state K, that is, the Fermi distribution function.

In order to evaluate the terms in square brackets, we set $n_\kappa=n_\kappa^0-\Delta n_\kappa$ where n_κ^0 is the value of n_κ when the spin-wave system is in equilibrium with the s-electrons and Δn_κ is the excess population in the state κ due to the resonance absorption. If $\Delta n_\kappa=0$, the terms

[12] T. Holstein and H. Primakoff, Phys. Rev. 58, 1098 (1940).
[13] C. Herring and C. Kittel, Phys. Rev. 81, 869 (1951).
[14] D. Bohm and D. Pines, Phys. Rev. 82, 625 (1951).

in square brackets vanish if one accounts for the energy δ function. Therefore, $\dot{n}_\kappa$ is proportional to Δn_κ and the terms in question become

$$-(df_K/d\epsilon_K)\epsilon_\kappa\Delta n_\kappa. \tag{10}$$

Here we have taken into account the energy δ functions of Eq. (9) and expanded the Fermi factors f near the top of the Fermi distribution.[15] The procedure is justified here since the maximum spin wave energy is of order one-fortieth of the Fermi energy.

The resulting expression for $\dot{n}_\kappa$ in which (10) replaces the square bracket of Eq. (9) is treated in the following way: The sum is transformed into an integral over $\mathbf{K}$ space with the introduction of a factor $2/(2\pi)^3$ where the 2 in the numerator accounts for two possible spin directions. The energy δ function is expressed in terms of $\mathbf{K}$, κ and the angle between them, θ, by the substitution $\mathbf{K}'=\mathbf{K}-\kappa$.

$$\delta(\epsilon_\kappa+\epsilon_{K'}-\epsilon_K)=(m/\hbar^2K\kappa)\delta[\cos\theta-F(\kappa,K)],$$
$$F(\kappa,K)=[(\hbar^2\kappa^2/2m)+\epsilon_\kappa]/(\hbar^2K\kappa/m). \tag{11}$$

The δ function is annihilated by the integration over θ and the remaining angular integration is trivial. The result is

$$\dot{n}_\kappa=-(16\pi e^2g\beta M_s/3mc^2\hbar)\Delta n_\kappa$$
$$\times[1-(q/\kappa)\tan^{-1}(\kappa/q)]^2(\epsilon_\kappa/\kappa^3)$$
$$\times\int K^3dK(df_K/d\epsilon_K)[1-F^2(\kappa,K)],$$
$$F^2(\kappa,K)\leq 1. \tag{12}$$

The condition $F^2\leq 1$ arises from the integration over θ where we must have $\cos^2\theta\leq 1$. The factor $df_K/d\epsilon_K$ behaves as a δ function, $\delta(\epsilon_K-E_F)$ where E_F is the Fermi energy for the s-electrons. This and the fact that the wave numbers of interest for the spin waves are of order 10^5 assure the smallness of F and it will henceforth be neglected. The integral is easily evaluated with this δ-function approximation. The result is

$$1/\tau=-\dot{n}_\kappa/\Delta n_\kappa=(32\pi me^2g\beta M_s/3c^2\hbar^5)$$
$$\times[1-(q/\kappa)\tan^{-1}(\kappa/q)]^2(\epsilon_\kappa/\kappa^3)E_F. \tag{13}$$

This result is the relaxation time for spin waves of wave number κ. In the case at hand we set $\kappa=1/\delta=(4\pi\omega\mu_2/\rho c^2)^{\frac{1}{2}}$ and find the relaxation time for those spin waves which are excited by the microwave field. Here ρ is the resistivity, ω is the angular frequency of the microwaves, and $i\mu_2$ is the permeability at resonance. For nickel, at 24 400 mc/sec, we have $\mu_2=17$,[5] $\rho=7.74$ micro-ohm cm,[16] and $\delta=1.5\times10^{-5}$ cm at room temperature. Thus, from Eq. (13), we find the relaxation time at room temperature in nickel

$$\tau=6\times10^{-8}\text{ sec},\quad T=293°\text{K}. \tag{14}$$

[15] The method used here is similar to that used in another connection by A. W. Overhauser, Phys. Rev. 89. 689 (1953).

The temperature dependence below room temperature is quite insignificant. The reason for this is that as the temperature is lowered, the skin depth is reduced in magnitude as the resistivity decreases. This results in a decrease in the factor $\kappa^{-3}=\delta^3$ of Eq. (13) which is offset by a concomitant increase of the screening factor which appears in the square brackets. The experimental value of Jan and Gijsman[16] for the resistivity at 14°K is about 1/20 of the room temperature result. If we use this value in our formulas we obtain

$$\tau=2\times10^{-8}\text{ sec},\quad T=14°\text{K}. \tag{15}$$

One would ordinarily expect a somewhat stronger temperature dependence of the resistivity in a very pure sample. However, even if we assume that the resistivity has decreased by as much as a factor 100, the relaxation time at 14°K remains of the same order because of the compensation of the aforementioned screening and skin depth effects. At room temperature and above, the relaxation time increases with δ because of the increased effectiveness of the screening, although we must realize that the spin-wave approximation is of doubtful validity before these temperatures are reached.

IV. INTERACTION WITH S-ELECTRON SPINS

The perturbation operator for the interaction of spin waves with s-electron spins is given by

$$\mathcal{H}_2=\int\begin{pmatrix}\varphi\\ \psi\end{pmatrix}^*\beta\boldsymbol{\sigma}\cdot\mathbf{H}_s\begin{pmatrix}\varphi\\ \psi\end{pmatrix}d\mathbf{r},$$

where[13] $\mathbf{H}_s=-2\pi(2g\beta M_s)^{\frac{1}{2}}(\boldsymbol{\kappa}/\kappa^2)(\alpha_\kappa\kappa^-+\alpha_{-\kappa}{}^*\kappa^+)e^{i\boldsymbol{\kappa}\cdot\mathbf{r}}$. Evaluation of this expression gives

$$\mathcal{H}_2=-2\pi\beta(2g\beta M_s)^{\frac{1}{2}}\sum_K\kappa(1/\kappa^2)(\alpha_\kappa\kappa^-+\alpha_{-\kappa}{}^*\kappa^+)$$
$$\times[b_K{}^*d_{K'}\kappa^-+d_K{}^*b_{K'}\kappa^+-(b_K{}^*b_{K'}-d_K{}^*d_{K'})\kappa_z]$$
$$\times\delta(\boldsymbol{\kappa}+\mathbf{K}'-\mathbf{K}). \tag{16}$$

The first two terms in square brackets involve s-electron spin flips while the third term does not. Consider first that process in which an s-electron flips from down to up and a spin wave is destroyed. This and its inverse process together give the collision rate by an expression exactly similar to Eq. (8) where, for example,

$$H_2{}'=-2\pi\beta(2g\beta M_s)^{\frac{1}{2}}(\kappa^+\kappa^-/\kappa^2)$$
$$\times[(n_\kappa+1)(1-f_{K'}{}^-)f_K{}^+]^{\frac{1}{2}}\delta(\boldsymbol{\kappa}+\mathbf{K}'-\mathbf{K}).$$

Here we have taken the matrix elements of the creation and destruction operators. $f_K{}^+$, for example, denotes the Fermi factor for wave vector $\mathbf{K}$ and spin up. For this interaction, the rate of energy transfer is given by

$$\dot{n}_\kappa=(128\pi^3\beta^3gM_s/15\hbar)\sum_K\kappa\epsilon_\kappa[(n_\kappa+1)(1-f_{K'}{}^-)f_K{}^+$$
$$-n_\kappa(1-f_K{}^+)f_{K'}{}^-]\delta(\epsilon_{K'}{}^-+\epsilon_\kappa-\epsilon_K{}^+), \tag{17}$$
$$\boldsymbol{\kappa}+\mathbf{K}'=\mathbf{K},$$

where we have averaged over directions of $\boldsymbol{\kappa}$. Here

[16] J. P. Jan and J. M. Gijsman, Physica 18, 339 (1952).

$\epsilon_K^{\pm}=(\hbar^2K^2/2m)\pm\beta H$. The terms in square brackets are handled in a way similar to the treatment of Eq. (9). We assume that there are equal numbers of s-electron spins up and down with small error since $\beta H/kT\ll 1$. With this assumption, as in the case of Eq. (9), the terms in square brackets reduce to Eq. (10) with the same notation.

We now proceed as before. The energy δ function becomes $(m/\hbar^2K\kappa)\delta(\cos\theta-F^-)$, where F^- is as in Eq. (11) save that $-2\beta H$ is added to the numerator. If we had considered instead the process in which the electron of wave vector $\mathbf{K}$ has down-spin and $\mathbf{K}'$ up-spin rather than the reverse, we would have gotten a factor F^+ with $+2\beta H$ added to the numerator. A detailed examination of the resulting expression for $\dot{Q}$ shows that we may neglect the distinction between F, F^+, and F^- since $\beta H\ll E_F$. We therefore account for the two kinds of processes mentioned here by the replacement of $F^{\pm}$ by F and the introduction of a factor 2. From Eq. (17) we obtain

$$\dot{n}_\kappa=-(64\pi m\beta^3gM_s/15\hbar^3)\Delta n_\kappa(\epsilon_\kappa/\kappa)\times\int KdK(df_K/d\epsilon_K),\quad F^2(\kappa,K)\leq 1. \quad (18)$$

As was the case with Eq. (12), $F^2\leq 1$ for κ near $1/\delta$ and K at the top of the Fermi distribution. The integral then reduces immediately to $(m/\hbar^2)$.

If we had considered the third term in square brackets of Eq. (16), i.e., processes in which the s-electron spin does not flip, then we would have gotten the same expression as Eq. (18) but reduced by a factor 1/4. To take all processes into account we need only multiply Eq. (18) by 5/4. The relaxation time for the spin-wave —s-electron interaction is therefore given by

$$1/\tau=-\dot{n}_\kappa/\Delta n_\kappa=(8\pi/3)(m^2g\beta^3M_s/\hbar^5)(\epsilon_\kappa/\kappa). \quad (19)$$

We again set $\kappa=1/\delta$ and find for nickel,

$$\tau=2\times10^{-9}/\delta=10^{-4}\text{ sec at }T=293^\circ\text{K}$$
$$=10^{-3}\text{ sec at }T=14^\circ\text{K}.$$

V. CONCLUSION

The interaction considered in Sec. IV is negligible compared with the "spin-current" interaction of Sec. III. The latter results in a relaxation time whose temperature dependence does not contradict the experimental results for the line width between liquid hydrogen and room temperatures. However, the relaxation time appears to be at least one order of magnitude too large to account for the magnitudes of the observed widths.

The present calculation does not apply to low-resistivity ferrites[17] because the free electron model cannot work well for d-band semiconductors such as the ferrites.

It is of interest to discuss the relaxation time for the "spin-current" interaction in the case that interactions between ferromagnetic spins are so strong that they account for the line width. In this case, according to KA, the spin system is in an equilibrium state before other relaxation processes have progressed appreciably. The relaxation time for energy transfer to the lattice via the conduction electrons is then determined by the spin-spin relaxation time (line width) itself for the following reason: the spin-current relaxation time for spin waves having wave numbers near q is very short indeed due to the negligible screening and the smallness of q (10^5 cm^{-1}). In this connection see Eq. (13). The rate of energy transfer to the lattice is then limited by the rate at which energy can be transferred from other spin-wave states to those spin waves with wave numbers near q by means of spin-spin interactions. This is because the relaxation primarily takes place through the spin-wave states of wave number q and at a rate somewhat faster than the spin-spin relaxation time.

It should be noted that the calculation begins to lose validity because of the inadequacy of the spin-wave approximation as the temperature rises above about one tenth of the Curie temperature (631°K for nickel). At higher temperatures, one would expect that the relaxation times will become somewhat shorter than those predicted here. Within the limits of the spin-wave approximation, however, the interactions discussed here do not appear to be strong enough to account for the experimentally observed line widths, which correspond to relaxation times of $10^{-9}-10^{-10}$ sec. The interaction of spin waves and conduction electron currents is stronger, however, than the direct spin-lattice interaction calculated previously,[6] especially at low temperatures where it gives much the shorter relaxation time.

The author would like to thank Professor C. Kittel for suggesting this problem and Dr. Y. Yafet for interesting discussion. He expresses his gratitude to Dr. G. T. Rado for reading the manuscript and making instructive comments regarding the importance of the effects arising from the finite skin depth.

Note added in proof.—A paper of T. Kasuya [Busseiron Kenkyu **74**, 1 (1954)], in which a calculation is made which is similar to the one reported here has come to the author's attention. Kasuya, however, fails to include the effect of screening of the magnetic field of the s-electron currents; nor does he attribute the breakdown to the $\kappa=0$ selection rule to the effect of the finite skin depth. The neglect of these effects yields a relaxation time an order of magnitude shorter than the result of the present work.

[17] Galt, Yager, and Merritt, Phys. Rev. **93**, 1119 (1954).

Theory of Ionic Ordering, Crystal Distortion, and Magnetic Exchange Due to Covalent Forces in Spinels*

J. B. GOODENOUGH AND A. L. LOEB
Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts
(Received September 1, 1954)

Elastic- and electrostatic-energy considerations are insufficient for an explanation of cation ordering between tetrahedral and octahedral sites in spinels. Tetrahedral, octahedral, and square covalent bonds are also important when cations are to be accommodated in these sites. The square bonds in octahedral sites can cause the tetragonal distortion observed in Mn_3O_4, γ-Mn_2O_3, $ZnMn_2O_4$, $CuFe_2O_4$, $CuCr_2O_4$, $CaIn_2O_4$, $CdIn_2O_4$, and metallic indium. A new magnetic exchange mechanism, "semicovalent exchange," which is consistent with the covalent model, is used to explain the magnetic properties of spinels.

I. INTRODUCTION

EXPERIMENTAL data on spinels indicate that the distribution of cations over tetrahedral and octahedral sites cannot be explained on the basis of electrostatic and elastic energies alone. Electronic bonding energies must be considered. Also, no explanation has yet been given for the difference beween Fe_3O_4, which alone transforms on cooling from a cubic to an orthorhombic spinel at a relatively low transition temperature, and other spinels, such as Mn_3O_4, $ZnMn_2O_4$, $CuFe_2O_4$, which become on cooling tetragonal at a relatively high transition temperature. The tetragonality of metallic indium and of $CaIn_2O_4$ and $CdIn_2O_4$ have never been explained. For hausmannite there has been no definitive selection between the formulas $MnO \cdot Mn_2O_3$ and $2MnO \cdot MnO_2$. The difference in resistivity discontinuities through the transition temperatures of Fe_3O_4, on the one hand, and Mn_3O_4 and $ZnMn_2O_4$, on the other, has not been explained. Neither has any explanation been given for the fact that Mn_3O_4 is paramagnetic at least down to 72°K despite the large magnetic moments of Mn^{2+}, Mn^{3+}, and Mn^{4+} ions. In this paper it is suggested that covalent forces are responsible for the above effects. A new magnetic-exchange mechanism, "semicovalent exchange," is proposed.

II. IONIC ORDERING

A. Description of Spinels

Many oxides consist of a face-centered cubic lattice of O^{2-} ions which is held together by interstitial cations. There are two kinds of interstices between the elements of such a lattice, tetrahedral interstices with four nearest neighbors and octahedral interstices with six nearest neighbors. If all the tetrahedral sites are empty and all the octahedral sites are filled, the crystal has the rock-salt structure; if only tetrahedral sites are occupied, the crystal has the zinc-blende structure. In oxides with the spinel structure there are cations in each type of interstice; twice as many octahedral as tetrahedral sites are occupied. The spinel lattice may be thought of as an ordered mixture of the zinc-blende and rock-salt structure.

There are two general classes which have the spinel lattice, the 2–4 spinels with the formula $2X^{2+}O \cdot Z^{4+}O_2$ and the 2–3 spinels $X^{2+}O \cdot Y_2^{3+}O_3$. In order to describe the cation distribution, these formulas may be written as $Z(X_2)O_4$ or $X(ZX)O_4$ and $X(Y_2)O_4$ or $Y(XY)O_4$, where the ions within the parentheses occupy octahedral sites. If there is only one kind of cation in the octahedral sites, the spinel is called normal. If there are equal amounts of both kinds of cations in the octahedral sites, the spinel is inverse. If there is an unequal number of each kind of cation in the octahedral sites, the spinel is mixed.

B. Critique of Previous Work

Since the physical properties of any spinel depend not only on the kind of cations in the spinel but also on their distribution within the interstices, it is important to understand what forces are responsible for cation ordering within a lattice. Verwey and his co-workers[1–14] have made a study of the cation arrangement in oxides with spinel structure. A summary of the experimental determination of cation location is given in Table I.

Since the observed cation distributions do not always

* The research in this document was supported jointly by the Army, Navy, and Air Force under contract with the Massachusetts Institute of Technology.

[1] E. J. W. Verwey, Z. Krist. **91A**, 65 (1935).
[2] van Arkel, Verwey, and van Bruggen, Rec. trav. chim. **55**, 331 (1936).
[3] Verwey, van Arkel, and van Bruggen, Rec. trav. chim. **55**, 340 (1936).
[4] E. J. W. Verwey and J. H. de Boer, Rec. trav. chim. **55**, 531 (1936).
[5] J. H. de Boer and E. J. W. Verwey, Proc. Phys. Soc. (London) **49A**, E59 (1937).
[6] Verwey, Haayman, and Heilmann, Philips Tech. Rev. **9**, 185 (1947).
[7] E. J. W. Verwey and E. L. Heilmann, J. Chem. Phys. **15**, 174 (1947).
[8] Verwey, Haayman, and Romeijn, J. Chem. Phys. **15**, 181 (1947).
[9] Verwey, de Boer, and van Santen, J. Chem. Phys. **16**, 1091 (1948).
[10] de Boer, van Santen, and Verwey, J. Chem. Phys. **18**, 1032 (1950).
[11] J. H. van Santen, Philips Research Repts. **5**, 282 (1950).
[12] E. W. Gorter, Compt. rend. **230**, 192 (1950); Nature **165**, 798 (1950).
[13] E. J. W. Verwey *et al.*, Z. physik. Chem. **198**, 6 (1951).
[14] F. C. Romeijn, Philips Research Repts. **8**, 304 (1953).

TABLE I. Observed cation distributions in some oxides with the spinel structure.[a]

X^{2+}	Y^{3+} or Z^{4+} Ionic radius[b]	Al^{3+} 0.57	V^{3+} 0.66	Cr^{3+} 0.64	(Mn^{3+}) 0.70	Fe^{2+} 0.67	Rh^{3+} 0.69	Ga^{3+} 0.62	In^{3+} 0.92	Ge^{4+} 0.44	Sn^{4+} 0.74	Ti^{4+} 0.64	Si^{4+} 0.39	V^{4+} 0.61
Mg^{2+}	0.78	N[c]	N[d]	N[c]		I†[f]	N[g]	I†[h]	I†[f]	O[e]		I[c]	O[i]	I[g]
Mn^{2+}	0.91	N[c]	N[d]	N[c]	(N)[c]	I[c]								
Fe^{2+}	0.83	N[c]	N[d]	N[c]		I[c]								
Co^{2+}	0.82	N[c]		N[c]		I[c]				N[c]	I[c]	I[c]		
Ni^{2+}	0.78	$\frac{1}{4}N\frac{3}{4}I$[c]		N[c]		I[c]				N[c]				
Cu^{2+}	(0.70?)	$[N]$[j]		I[k]		I†[c]								
Zn^{2+}	0.83	N[c]	N[d]	N[g]	(N)[c]	N[c]	N[g]	$[N]$[l]			I[f]	I[c]	P[i]	
Cd^{2+}	1.03	N[c]		N[c]		N[c]		N[g]	$[N]$[j]					
Li^{+}	0.78	$Al(Li_{\frac{1}{4}}Al_{\frac{3}{4}})O_4$[c]		$Fe_{\frac{1}{4}}Li_{\frac{1}{4}}(Cr_2)O_4$[c]				$Fe(Li_{\frac{1}{4}}Fe_{\frac{3}{4}})O_4$[c]						

[a] $[N]$ means normal structure inferred, cation distribution not experimentally established. Distributions marked with daggers are not entirely inverse; the number of divalent ions in the tetrahedral sites are temperature-dependent. The letters N, I, O, P refer to normal-spinel, inverse-spinel, olivine, and phenacite structures, respectively.

[b] *Atomic Radii*: *Handbook of Chemistry and Physics* (Chemical Rubber Publishing Company, 1953–1954), thirty-fifth edition. Goldschmidt values wherever available.

[c] F. C. Romeijn, Philips Research Repts. **8**, 304 (1953).

[d] G. H. B. Lovell, Trans. Brit. Ceram. Soc. **50**, 315 (1951).

[e] E. J. W. Verwey and E. L. Heilmann, J. Chem. Phys. **15**, 174 (1947).

[f] T. F. W. Barth and E. Posnjak, Z. Krist. **82**, 325 (1932).

[g] W. Rüdorf and B. Reuter, Z. anorg. Chem. **253**, 177, 194 (1947).

[h] F. Machatschki, Z. Krist. **82**, 348 (1932).

[i] A. F. Wells, *Structural Inorganic Chemistry* (Oxford University Press, London, 1950), second edition.

[j] L. Passerini, Gazz. chim. ital. **60**, 389, 754 (1930).

[k] McGuire, Howard, and Smart, Ceram. Age **60**, 22 (1952).

[l] F. Buschendorf, Z. physik. Chem. (B) **14**, 297 (1931).

coincide with their predictions, which were based only on the electrostatic and elastic forces in a purely ionic model, Verwey and Heilmann[7] suggested that the anomalies could be accounted for by the stability of the tetrahedral covalent bonds which form about certain cations when located in tetrahedral interstices.

Romeijn[14] has suggested that ions with a noble-gas outer-electron-shell structure, such as Mg^{2+}, Al^{3+}, and Ti^{4+}, are less compressible than ions with a full or half-filled d-shell, such as Zn^{2+}, Cd^{2+}, Ga^{3+}, In^{3+}, Ge^{4+}, Sn^{4+}, Mn^{2+}, and Fe^{3+}, so that the latter group is more easily accommodated in tetrahedral sites than is the former. It is not obvious why a full or half-filled d shell should be more compressible than a noble-gas shell. However, since the outer electron shells in noble-gas structures are separated from the next higher energy levels by a considerable energy gap, they do not provide empty orbitals for the formation of covalent bonds. In transition metals the nd and $(n+1)s$ orbitals differ very little in energy. Covalent bonds are therefore easily formed, and since covalent bonds involve orbital overlap, they are shorter than electrovalent bonds. Therefore the ions with outer d shell appear more compressible than those with outer noble-gas shell.

Finally Romeijn[14] and van Santen and van Wieringen[15] considered the effect on electronic binding energies of crystalline fields in ionic crystals of cubic symmetry. They stated that the d shell splits in such a field into two set of orbitals, one extending along the ⟨111⟩ directions, the other in the ⟨100⟩ directions.[16] Van Santen and van Wieringen assumed that the former is more stable when the outer electron configuration of the cation is d^3 or d^8 (as in Mn^{4+}, Cr^{3+}, and Ni^{2+}), but that the latter is more stable for the configuration d^2 or d^7 (as in V^{3+} and Co^{2+}). An ion with a stable ⟨111⟩ electron configuration would, according to van Santen and van Wieringen's express assumption of electrovalent bonding, fit more easily into an octahedral site where the electron charges of the cation can fit between the directed anion p orbitals without appreciable overlap. An ion with a stable ⟨100⟩ electron configuration would fit more easily into a tetrahedral site for an analogous reason.

The assumption regarding the relative stabilities of the d-shell bonds has never been substantiated. What is more important, however, is that the result of the overlapping of the filled oxygen p orbitals with the empty cation orbitals has been ignored. Such an overlap gives rise to coordinate covalence which reduces the polarity of the cation-anion bond, stabilizes the system, and reduces the effective anion radius. The influence of electrostatic and elastic energies on cation ordering is reduced. It will be shown below that if covalent considerations are used to supplement and modify the electrostatic and geometric considerations of the purely ionic model, the cation distributions in the spinel lattice can be readily accounted for without recourse to an assumption either with regard to the d-shell splitting of various atoms or to the compressibility of the d shell.

The cation distribution in a crystal is determined by its minimum free energy $A=E-TS$, where T is its absolute temperature, S is its entropy, and E, its internal energy, is the sum of an electrostatic contribution E_i, an elastic (Born-repulsion) contribution E_r, and an electronic-bonding and exchange contribution E_e. In the spinel lattice, covalence effects E_i, E_r, and E_e. In the absence of covalence E_e approximately equals zero. Covalent bonding makes E_e appreciably negative

[15] J. H. van Santen and J. S. van Wieringen, Rec. trav. chim. **71**, 420 (1952).

[16] H. A. Bethe, Ann. Physik **3**, 133 (1929).

and, by decreasing the ionic charge and bond length, alters the influence of E_i and E_r. Since disorder corresponds to high entropy and the entropy contribution (TS) increases with temperature, high temperatures tend to disturb ordering.

III. COVALENT BONDS

A. Introduction

It was pointed out in Sec. IIA, above, that spinels are an ordered mixture of the zinc-blende and rock-salt structures. It is expected that those metals whose oxides form the zinc-blende structure tend to occupy tetrahedral sites in spinels as they do in the oxides. Similarly metals whose oxides are of the rock-salt type tend to occupy octahedral sites. Frequently the relative stabilities of the two metals in a given interstice in the spinels must be compared before an actual prediction about their distribution over the two types of sites can be made.

The zinc-blende interionic bond is predominantly covalent if the cation is large. Purely ionic forces would place the cation in the interstices with the largest anion coordination, *viz.*, the octahedral position. The metals whose oxides have the zinc-blende structure are therefore expected to be covalently bound to tetrahedral sites when in a spinel lattice.

The octahedrally coordinated cations of the rock-salt structure may be either electrovalently or covalently bound to their near-neighbor anions. The relative strength with which a metal, whose simple oxide has the rock-salt structure, is held to an octahedral interstice in the spinel lattice depends upon the relative charge, size, and covalent-bonding of the competing cations. Cations which are more stable in octahedral sites in the simple oxides tend to go into octahedral sites, but may be forced into tetrahedral sites in the spinel lattice. Therefore both electrovalence and covalence may be found in each type of interstice. To distinguish between electrovalence and covalence, the valence of the former is written in Arabic numerals (e.g., Mg^{2+}), of the latter, in Roman numerals (e.g., Zn^{II}).

In spinel crystals the coordination of the close-packed-anion interstices limits the possible cation-anion bond types which can be formed. If the stable covalent-bond types which any cation forms can be independently determined, the anion interstices in which the steric constraints are compatible with stable covalent-bond formation are known. Measurements of the steric structure of complex ions containing a particular cation provide the necessary independent information since the complex-ion structure is primarily determined by the stable covalent-bond formation. Complex-ion structure is only secondarily effected by steric hindrance and the difference in the electronegativities of the components. The steric structure of complex ions can be determined by optical and by magnetic-susceptibility measurements.[17] Thus a knowledge of complex-ion structure gives information about both the relative stabilities of a cation in octahedral *vs* tetrahedral coordination and the possible distortion of the steric constraints.

If covalent bonds are formed in the spinel lattice, both bonding electrons come from the anion, and the bond is coordinate-covalent. The near-neighbor anions perturb the cations so that they form stable, empty hybrid orbitals which overlap with the anion p orbitals. If the empty orbitals have energies which are nearly degenerate with the outermost filled orbitals, if the electronegativity difference between the ions is not too great, and if the orbital overlap is large, the empty cation orbitals interact with the full anion p orbitals to accept electrons from the anions. Then the anion p electrons are induced to share their time between the cation and the anion. If empty orbitals are not available for hybridization, half-filled orbitals can participate in covalence. However, filled cation orbitals obviously do not participate in covalent-bond formation.

B. The Oxygen Orbitals

The oxygen ion has the structure $1s^2 2s^2 2p^6$. The p orbitals are three mutually perpendicular dumbbells. In oxides with the spinel lattice each oxygen has four near-neighbor cations, of which three are in octahedral sites and the fourth is in a tetrahedral site. The three octahedral sites are situated along three mutually perpendicular directions from the oxygen; the tetrahedral site lies in the $\langle 111 \rangle$ direction away from the quadrant defined by the three octahedral sites. The oxygen can bond covalently with any or all of the octahedral-cation neighbors through its three p orbitals. In this case, however, there are no anion orbitals directed toward the tetrahedral site, and the tetrahedral-site cation-oxygen bond is ionic.

If the bond toward the tetrahedral site is covalent, the overlap of full oxygen orbitals with the empty orbitals of the octahedral-site cations is reduced since the oxygen p orbitals no longer point directly towards the cations. This reduction in orbital overlap reduces the strength of any covalent bonding in the octahedral sites.

Thus there is competition between the tetrahedral and octahedral sites for overlap with the full oxygen p orbitals. Unless strong covalence can occur simultaneously in both types of sites, there will be covalent bonding in, at most, one type of site only. If stable covalent bonds can be formed in both types of sites at once, hybrid (sp^3) orbitals, which point toward the four corners of a tetrahedron, may be admixed with the p^3 orbitals on the oxygen ion. Such orbitals do not overlap as much with the octahedral-cation orbitals as do the pure p^3 orbitals, and the strength of the octahedral-site

[17] L. Pauling, *The Nature of the Chemical Bond* (Oxford University Press, London, 1948), second edition; J. Am. Chem. Soc. 53, 1367 (1931).

covalent bond is reduced from the value it has when the tetrahedral-site bonds are ionic. A semiquantitative estimate of the proportionate reduction in octahedral-site covalent-bond strength due to simultaneous covalence in the tetrahedral sites is given later in Appendix I and Fig. 6.

C. Cation Hybrid Orbitals

There are three kinds of cation hybrid orbitals which are important in the spinel lattice where the steric constraints are due to the close packing about either tetrahedral or octahedral interstices. If either steric hindrance or too large differences in electronegativities prevent hybrid covalent-bond formation, ionic bonds are formed. The important cation hybrid orbitals have tetrahedral, octahedral, or square configurations.

1. Tetrahedral Orbitals

The tetrahedral orbitals result from the hybridization of (sp^3) orbitals.[17] These point toward the four corners of a tetrahedron. They are predominantly formed in cations with a full d shell which is separated in energy by only a small interval from the next empty s and p shells. Tetrahedral coordination is common for Zn^{II} and Cd^{II}, for instance. The strength of these covalent bonds decreases with increasing atomic number of the cation and decreasing atomic number of the anion in any one column of the periodic table.

There is some evidence that Fe^{3+} and Mn^{2+}, each with a half-full, spherically symmetric d shell, form bonds which are partially covalent. Although these ions form ionic bonds in Fe_2O_3 and MnO, the wurtzite structure of one phase of MnS indicates that the possibility of covalent-bond formation is by no means negligible. In the tetrahedral sites of the spinel lattice, where the steric constraints favor the formation of tetrahedral covalent bonds, these ions undoubtedly form bonds with a partially covalent character. Evidence for this lies not only in the fact that the ferrospinels are frequently inverse, with a definite tendency for Fe^{III} ions to go into the tetrahedral sites, but also in the observation that inverse ferrospinels have a lattice parameter which is roughly 0.06 A smaller than would be expected if the lattice were normal.[7] This reduction in lattice parameter could be due to a shorter cation-anion bond for Fe^{III} in the tetrahedral sites than Fe^{3+} in the octahedral sites, the difference in bond length indicating the degree of covalence in the tetrahedral bonds.

2. Octahedral Orbitals

The octahedral, hybrid (d^2sp^3) orbitals occur in the transition elements which have partially filled d shells. In this case six bonds are directed toward the centers of six faces of a cube. These bonds are not easily identified since they are accommodated, without distortion, in an octahedral site just as are ions which form predominantly ionic bonds. These bonds would form most readily about cations with an outer configuration d^3 or d^6. If an ion with electron configuration d^6 forms an ionic bond in an octahedral site, it contributes 4.90 Bohr magnetons to the magnetic susceptibility whereas it contributes no paramagnetic moment if it is bound covalently.[17] The bond cannot be partially ionic and partially covalent since this would require mixing two states of different electron spin. In complex ions like $[Fe(CN)_6]^{4-}$ the Fe^{II} forms covalent bonds, but in the oxides the large electronegativity of the oxygen ions causes Fe^{2+} to form ionic bonds. In d^3 ions like Cr^{III} and V^{II}, however, there is no spin change involved in changing from ionic to covalent bonding so that partially covalent bonds can be formed which cannot be distinguished from ionic bonds by magnetic measurements. It is probable that these ions are stabilized in the octahedral sites by partial covalent-bond formation.

3. Square Orbitals

Finally there are the square, coplanar (dsp^2) orbitals which point toward the four corners of a square. These occur in the transition elements that have one d orbital available for hybridization, such as cations with outer electron configuration d^4, d^8, or d^9. Cations which form square, covalent bonds can be accommodated in an octahedral site.

The cation Cu^{II} is known to form square bonds in complex ions,[18] cupric oxide,[19] and cupric salts.[19] Although the Cu^{II} cation may, theoretically, form either four tetrahedral (sp^3) orbitals or four square (dsp^2) orbitals, the square-bond formation in CuO clearly indicates that the latter bond type is the more stable bond for Cu^{II} in an oxide lattice. Therefore the Cu^{II} cation is more stable in the octahedral sites of the spinel lattice where the steric constraints favor the formation of square bonds.

Since cations with the outer electron configuration d^8 can accommodate at most only ten additional electrons in a hybrid dsp shell, their maximum coordination number is five. Therefore they cannot form octahedral (d^2sp^3) bonds since this would require exciting two electrons to some higher energy level. Since there are two vacancies in the d shell, square (dsp^2) orbitals can be formed. In order to hybridize such orbitals, however, the cation must reverse one electron spin to reduce its paramagnetic moment to zero. There are many nickel, palladium, and platinum complexes,[17,19] such as $[Ni(CN)_4]^{2-}$, which are both coplanar and diamagnetic. It is significant that in their cyanide complexes Ni^{II} has coordination number four whereas Fe^{II} and Fe^{III} have coordination number six. Because an electron spin is reversed when d^8 cations form covalent bonds, the bonds these cations form are either covalent or ionic, but never a mixture of the two. Since NiO has the rock-salt

[18] D. P. Mellor and D. P. Craig, Proc. Roy. Soc. (N. S. Wales) 75, 27 (1941).

[19] A. F. Wells, *Structural Inorganic Chemistry* (Oxford University Press, London, 1950), second edition.

structure with the Ni^{2+} electrovalently bonded, it may be assumed that Ni^{2+} forms electrovalent bonds in the oxides with the spinel structure. The large difference in electronegativities between nickel and oxygen is the decisive factor favoring ionic-bond formation. Palladium, however, is less electropositive than nickel and does form square bonds in PdO. If Pd^{II} were in a spinel, it should form square bonds in the octahedral sites.

Cations with the outer electron configuration d^4, such as Mn^{3+}, can hybridize empty (dsp^2) orbitals for coordinate covalence without altering their paramagnetic moment. Since it would require energy to reverse an electron spin to form (d^2sp^3) orbitals, the square bond is the most stable covalent-bond configuration. These ions can be accommodated in octahedral sites with the formation of partially or completely covalent bonds.

D. Summary

In Table II there is a summary of the covalent-bond strengths and the cation distributions favored by the covalent forces. The strength estimates depend, of course, upon the relative sizes of the cations and the interstices. They also depend upon the electronegativity difference between the anions and the cations. Finally, they depend upon the strength of the covalent bond in the opposite set of lattice sites. If a bond cannot be partially covalent, covalent effects have no influence unless the bond is completely covalent. If such a bond is marked "weak" in Table II it means it has a small chance of being completely covalent. If such a covalent bond is formed, however, its influence is strong.

As a rule of thumb, it is suggested that the covalent forces be considered first in any prediction of the cation distribution in a spinel lattice. If these are weak, the energy term E_e may be neglected and the Madelung energy and geometrical considerations of Verwey and his co-workers may be used. If the covalent forces are moderate or strong, these are usually the predominant ordering influence. With such a simple, qualitative rule, all of the observed structures given in Table I, except $Co(CoSn)O_4$, can be predicted. Since the Madelung energy strongly favors the observed structure in $Co(CoSn)O_4$, it is concluded that the covalent forces, which are weaker in this instance than when associated with any of the other Group IV cations, are not sufficiently strong to reverse the order.

Mg_2GeO_4 and $MgSiO_4$ have the olivine structure. The olivine structure may be described either as an assembly of $[GeO_4]^{4-}$ or $[SiO_4]^{4-}$ groups with the Mg^{2+} ions between them, or as an infinite array of approximately close-packed oxygen ions with Ge^{IV} or Si^{IV} in tetrahedral and Mg^{2+} in octahedral interstices. This structure is analogous to the spinel. It differs, however, in that the tetrahedral bonding is so much stronger than the octahedral bonds that the cubic close packing of the oxygen ions is distorted.

The phenacite structure, which is formed by Zn_2SiO_4, has a tetrahedral coordination about both the Zn^{II} and the Si^{IV} cations. The Si—O and Zn—O covalent bonds are apparently sufficiently strong to force all the cations into tetrahedral sites and distort the oxygen lattice from close packing.

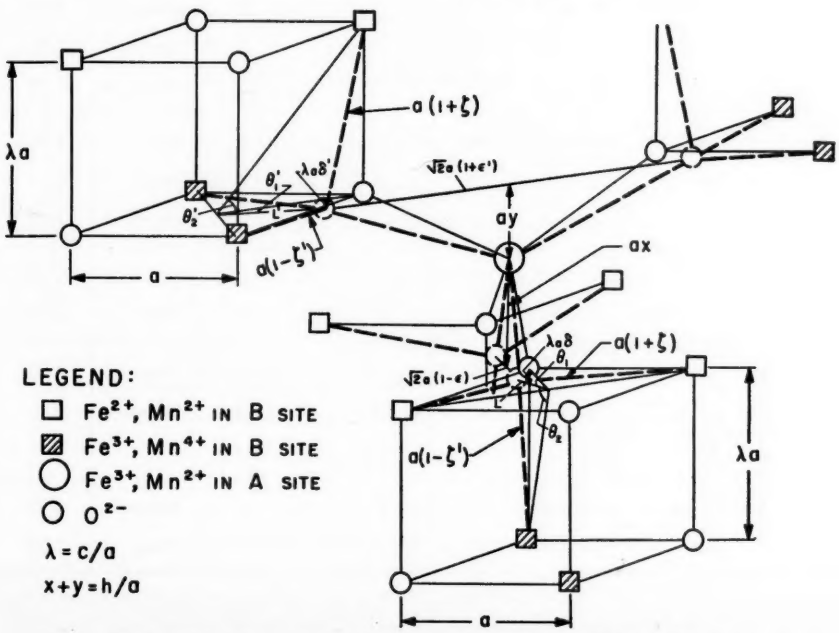


FIG. 1. Ionic model of distortion to tetragonal symmetry due to ordering of *B*-site ions.

TABLE II. Estimated strengths for possible covalent forces in spinel lattices and predicted magnetic moments in Bohr magnetons for cations bonded in various ways.

A. Transition cations.

Config.	Cations	Orbitals	Electron-spin config.[a] d	s	p	Estimated strength[b]	Magnetic moment
d^0	Ti^{IV} V^{V} Cr^{VI}	Tet. (sp^3)	⇅	⇅	⇅	Moderate	0
	Zr^{IV} Nb^{V} Mo^{VI}	Oct. (d^2sp^3)	⇅ ⇅	⇅	⇅	Very weak	0
	Hf^{IV} Ta^{V} W^{VI}	Sq. (dsp^2)	⇅	⇅	⇅	Very weak	0
d^1	V^{IV}	Tet. (sp^3)	→	⇅	⇅	Weak	1.73
	Nb^{IV} Mo^{V}	Oct. (d^2sp^3)	→ ⇅ ⇅	⇅	⇅	Very weak	1.73
	W^{V}	Sq. (dsp^2)	→ ⇅	⇅	⇅	Very weak	1.73
d^2	V^{III} Cr^{IV}	Tet. (sp^3)	→ →	⇅	⇅	Very weak	2.83
	Mo^{IV} Ru^{VI}	Oct. (d^2sp^3)	→ → ⇅ ⇅	⇅	⇅	Weak	2.83
	W^{IV} Os^{VI}	Sq. (dsp^2)	→ → ⇅	⇅	⇅	Very weak	2.83
d^3	V^{II} Cr^{III} Mn^{IV}	Tet. (sp^3)	→ → →	⇅	⇅	Very weak	3.88
	Mo^{III}	Oct. (d^2sp^3)	→ → → ⇅ ⇅	⇅	⇅	Moderate	3.88
		Sq. (dsp^2)	→ → → ⇅	⇅	⇅	Weak	3.88
d^4	Cr^{II} Mn^{III} Fe^{IV}	Tet. (sp^3)	→ → → →	⇅	⇅	Weak	4.90
	Mo^{II} Ru^{IV}	Oct. (d^2sp^3)	⇄ → → ⇅ ⇅	⇅	⇅	Very weak	2.83
	Os^{IV}	Sq. (dsp^2)	→ → → → ⇅	⇅	⇅	Strong	4.90
d^5	Mn^{II} Fe^{III} Co^{IV}	Tet. (sp^3)	→ → → → →	⇅	⇅	Moderate	5.91
	Ru^{III}	Oct. (d^2sp^3)	⇄ ⇄ → ⇅ ⇅	⇅	⇅	Very weak	1.73
	Os^{III} Ir^{IV}	Sq. (dsp^2)	⇄ → → → ⇅	⇅	⇅	Very weak	3.88
d^6	Fe^{II} Co^{III}	Tet. (sp^3)	⇄ → → → →	⇅	⇅	Weak	4.90
	Ru^{II} Rh^{III} Pd^{IV}	Oct. (d^2sp^3)	⇄ ⇄ ⇄ ⇅ ⇅	⇅	⇅	Weak	0
	Ir^{III} Pt^{IV}	Sq. (dsp^2)	⇄ ⇄ → → ⇅	⇅	⇅	Very weak	2.83
d^7	Co^{II} Ni^{III}	Tet. (sp^3)	⇄ ⇄ → → →	⇅	⇅	Weak	3.88
	Rh^{II}	Sq. (dsp^2)	⇄ ⇄ ⇄ → ⇅	⇅	⇅	Weak	1.73
d^8	Ni^{II}	Tet. (sp^3)	⇄ ⇄ ⇄ → →	⇅	⇅	Weak	2.83
	Rh^{I} Pd^{II} Ag^{III}	Sq. (dsp^2)	⇄ ⇄ ⇄ ⇄ ⇅	⇅	⇅	Moderate	0
	Pt^{II} Au^{III}						
d^9	Cu^{II}	Tet. (sp^3)	⇄ ⇄ ⇄ ⇄ →	⇅	⇅	Moderate	1.73
		Sq. (dsp^2)	⇄ ⇄ ⇄ ⇄ ⇄	⇅	⇅	Strong	1.73

[a] Solid arrows refer to cation-electron spins; dashed arrows to spins of anion electrons entering into coordinate covalence.
[b] Strength estimates are modified by (1) relative size of interstice and cation (i.e., amount of orbital overlap), (2) difference in electronegativities of cation and anion, (3) type of bonding in other interstices, and (4) the atomic weight of the cation.

TABLE II (*continued*).

B. Other cations

Cation	Estimated strength[a]	Cation	Estimated strength[a]
Li^{+}	Electrovalent	Ga^{3+}	Mod. strong tet. (sp^3)
Mg^{2+}	Electrovalent	In^{3+}	Mod. tet. (sp^3) and sq. (dsp^2)
Al^{3+}	Electrovalent	(C^{4+})	Extremely strong tet. (sp^3)
Zn^{2+}	Strong tet. (sp^3)	Si^{4+}	Very strong tet. (sp^3)
Cd^{2+}	Strong tet. (sp^3)	Ge^{4+}	Strong tet. (sp^3)
Hg^{2+}	Mod tet. (sp^3)	Sn^{4+}	Mod tet. (sp^3)

[a] Strength estimates are modified by (1) relative size of interstice and cation (i.e., amount of orbital overlap), (2) difference in electronegativities of cation and anion, (3) type of bonding in other interstices, and (4) the atomic weight of the cation.

IV. LATTICE DISTORTION

It is proposed that in spinels the lattice distortions to tetragonal symmetry are due to covalent-bond formation whereas those to orthorhombic symmetry are due to cation ordering in the octahedral sites. In Sec. A, below, there is a discussion of the distortions due to ionic ordering. In Sec. B there is a theoretical discussion of tetragonal distortion due to coplanar covelant-bond formation and an application of the theory to $CuFe_2O_4$. In Sec. C both models are applied to spinels containing manganese. Experimental observations are shown to be consistent with the covalent model. In Sec. D it is proposed that the indates and metallic indium are also distorted by covalent-bond formation.

A. Lattice Distortion and Ionic Ordering

Since 1929[20,21] it has been known that magnetite (Fe_3O_4) undergoes a phase transition as the temperature is raised through 119°K. This phase transition produces a ninety-fold increase in the electrical conductivity over the temperature range $118.9°K \leq T \leq 119.9°K$.[8,22] After Verwey and de Boer[4] had determined that Fe_3O_4 is inverse, Verwey[23] suggested that below the transition temperature the Fe^{2+} and Fe^{3+} ions in the octahedral sites are ordered in alternating planes perpendicular to the c axis. Above the transition temperature disorder is produced by the movement of electrons between Fe^{2+} and Fe^{3+} ions. The transition temperature is low because the transition is electronic, rather than ionic.

If the lattice distortion below 120°K is due to cation ordering in an ionic lattice, it should be possible to calculate the ordered-phase lattice parameters from a knowledge of the ionic radii. Verwey predicted that a tetragonal distortion would result if there was an ordering of Fe^{2+} and Fe^{3+} cations into alternate planes of octahedral, or B, interstices. Such a distortion would be possible only in a purely ionic lattice in which the tetrahedral, or A, site is distorted as shown in Fig. 1. The generality of this distortion for all tetrahedral sites is apparent from the representation of the spinel structure in Fig. 2. Although, as was first suggested by Bickford,[24] orthorhombic symmetry will result if the tetrahedral site is not distorted as shown, there is no obvious reason why, if a purely ionic model is correct, a distortion to tetragonal symmetry should not occur provided the oxygen-oxygen separation is $\sqrt{2}a(1-\epsilon) > 2.64$ A.

The cation-oxygen distance depends upon the size of the cation. If the $Fe^{2+}-O^{2-}$ distance is $a(1+\zeta)$ and the $Fe^{3+}-O^{2-}$ distance is $a(1-\zeta')$, then the distortion

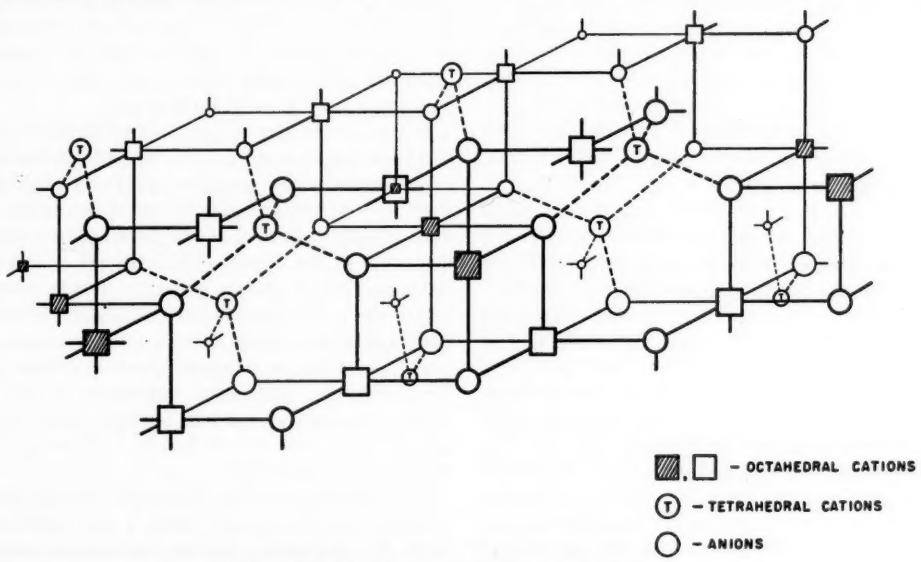


FIG. 2. The spinel lattice with ordering in the B sites.

[20] P. Weiss and R. Forrer, Ann. phys. (10) **12**, 279 (1929).
[21] R. W. Millar, J. Am. Chem. Soc. **51**, 215 (1929).
[22] B. A. Calhoun, Phys. Rev. **94**, 1577 (1954).
[23] E. J. W. Verwey, Nature **144**, 327 (1939); E. J. Verwey and P. W. Haayman, Physica **8**, 979 (1941).
[24] L. R. Bickford, Jr., Phys. Rev. **78**, 449 (1950).

of the tetrahedral interstice can be calculated from the geometrically derived relations:

$$\begin{aligned}(1+\epsilon)&=[2(1+\zeta)^2-1]^{\frac{1}{2}}\cos\theta_1,\\ \lambda\delta&=[(1+\zeta)^2-\tfrac{1}{2}]^{\frac{1}{2}}\sin\theta_1,\\ (1-\epsilon')&=[2(1-\zeta')^2-1]^{\frac{1}{2}}\cos\theta_1',\\ \lambda\delta'&=[(1-\zeta')^2-\tfrac{1}{2}]^{\frac{1}{2}}\sin\theta_1',\end{aligned}\qquad\text{(IV.1)}$$

where $\lambda a\delta$ and $\lambda a\delta'$ give the displacement of the oxygen ion from the planes of Fe^{2+} and Fe^{3+} ions, respectively; $\sqrt{2}a(1-\epsilon)$ and $\sqrt{2}a(1+\epsilon')$ give the separations of the tetrahedral oxygens near the Fe^{2+} and the Fe^{3+} planes, respectively; and θ_1, θ_1' are the declination angles shown in Fig. 1. The height $h=(x+y)a$ of the distorted tetrahedron can then be calculated from the geometric relation,

$$(h/a)=x+y=[d^2-\tfrac{1}{2}(1-\epsilon)^2]^{\frac{1}{2}}+[d^2-\tfrac{1}{2}(1+\epsilon')^2]^{\frac{1}{2}}. \quad \text{(IV.2)}$$

Since an ionic model is used, the tetrahedral-site $Fe^{3+}-O^{2-}$ separation, (ad), is assumed identical for each of the four oxygen ions. If the interionic distances are known, a self-consistent calculation for λ can be carried out provided use is made of the requirement,

$$(h/a)=\lambda(1+\delta-\delta'). \quad \text{(IV.3)}$$

If the values $a(1+\zeta)=2.15$ A, $a(1-\zeta')=2.09$ A, $a=2.10$ A, and $d=\sqrt{3}/2$ are chosen, then the predicted axial ratio for the tetragonal lattice is $c/a=\lambda=1.022$.[25]

Although tetragonal symmetry can be predicted by a purely ionic model in which a close packing of hard spheres is envisaged, orthorhombic symmetry must result if covalent forces play a significant role in the tetrahedral-site bonds. Since the covalent-bond contributions which are formed by overlap of Fe^{3+}-(sp^3) hybrid orbitals and the oxygen p wave functions have their maximum strength for tetrahedral coordination, they resist any distortion of the tetrahedral bond angles. The cation-anion bond angles in the distorted tetrahedral interstice of Fig. 1 would be more nearly tetrahedral if x were reduced and y were increased, i.e., if the cation were located closer to the neighboring B-site plane perpendicular to the c axis which contains Fe^{2+} cations than that which contains Fe^{3+} cations. Thus the resistance to distortion of the tetrahedral covalent-bond angles forces apart the B-site-Fe^{2+} cations and forces together the B-site-Fe^{3+} cations. Since the ordered, B-site-Fe^{2+} cations are arranged in parallel rows which are perpendicular to the similar set of parallel Fe^{3+} ions in alternate planes (see Fig. 2), a separation of the Fe^{2+} cations and an approach of the Fe^{3+} cations produces a shear strain so that an orthorhombic distortion results. The reported distortion is of this type. Abrahams and Calhoun[26] have measured with x-rays the dimensions of the unit cell at 78°K. They reported the orthorhombic axes to be $a=5.912$, $b=5.945$, $c=8.388$ A.

[25] See Appendix II.

[26] S. C. Abrahams and B. A. Calhoun, Acta Cryst. 6, 105 (1953).

This discussion leads to the generalization that if two kinds of cation occupy octahedral interstices and if they order in alternate planes perpendicular to some c axis, then the lattice will suffer a distortion to tetragonal or orthorhombic symmetry. The magnitude of the distortion will depend upon the difference in cation size, and the degree of orthorhombicity will increase with the strength of the covalent bonds in the tetrahedral interstices.

B. Lattice Distortion and Square-Bond Formation

1. Theory

If a cation hybridizes (dsp^2) orbitals to form four square, covalent bonds in an octahedral interstice of the spinel lattice, the steric constraints will permit covalent bonds to be formed in any one of the three biaxial planes of the site. At high temperatures the covalent bonds will resonate between the three possible biaxial planes. All six octahedrally directed bonds are partially ionic, partially covalent, and, since they are equivalent, the cubic symmetry of the interstice is preserved. Below some ordering temperature, however, the covalent bonds are frozen into one of the three biaxial planes, and the octahedral-site cation is bound to the six near-neighbor anions by four coplanar covalent bonds and two linear electrovalent bonds. Because of orbital overlap, the four coplanar covalent bonds are shorter than the two electrovalent bonds, and the elementary octahedron becomes distorted from cubic to tetragonal symmetry with an axial ratio $c/a>1$. If the elementary octahedra have tetragonal symmetry, the elastic energy of the crystal is minimized when their tetragonal axes are aligned parallel to one another. This produces a distortion of the bulk lattice from cubic to tetragonal symmetry with an axial ratio $c/a>1$.

If λ is the fraction of square-bond-forming cations in octahedral sites, there exists a value $\lambda=\lambda_0$ below which no distortion of the crystal occurs. For $\lambda<\lambda_0$ there is no alignment of tetragonally distorted octahedra to produce a bulk distortion of the lattice. For a tetragonal distortion of the crystal it is necessary to have a continuously linked system of square bonds. For such a long-range order to occur, each octahedral ion that does not form square bonds must, on the average, have at least two of its six octahedral near neighbors that do. Therefore, for tetragonal distortion of the crystal lattice to occur, at least 25 percent of the octahedral cations must form square bonds. This means that the minimum λ_0 equals 0.25.

If the band model is applied to the crystal, the highest occupied electron-energy band is the (dsp^2) covalent band. Because orbital overlap decreases with decreasing distortion, the energy of the (dsp^2) covalent band is raised with decreasing axial ratio so that it becomes easier to remove electrons from it. At high temperatures some of the electrons are thermally excited from the covalent band. However, when an electron is excited,

a hole is left in the covalent band so that the coplanar covalent bonds in some octahedral site become partially ionic. This ionic character resonates among three mutually perpendicular planes and reduces the difference between the character of the bonds parallel, and that of the bonds perpendicular, to the c axis so that the tetragonal distortion is reduced. This, in turn, facilitates the removal of an electron from the covalent band. Thus the loosening of the square bonds from the planes perpendicular to the c axis is a cooperative phenomenon, and the transition from the tetragonal to the cubic phase occurs over a small temperature interval when the sample is heated.

Conversely, if the sample is cooled from above the transition temperature, there is a temperature at which the covalent band is sufficiently full for some of the covalent bonds to become frozen into place. This "freezing in" of covalent bonds increases c/a and therefore the orbital overlap. An increase in the orbital overlap reduces the number of electrons which can be thermally excited from the covalent bond so that more covalent bonds are frozen in. Again there is a fairly sharp transition temperature, but, because of the difference in axial ratios, this temperature is lower than that observed when the sample is heated. This temperature difference shows up as a thermal hysteresis which varies in temperature change with the magnitude of the axial ratio of the tetragonal phase.

In an ordered-bond model there is a correlation between the lattice distortion and the excitation energy as measured by the temperature dependence of the conductivity. A conductivity discontinuity and thermal hysteresis at the transition temperature are correlated with the discontinuity and hysteresis in the axial ratio. Further, the excitation energy decreases slowly with increasing temperature so that a plot of the logarithm of the conductivity *vs* the reciprocal of the temperature does not give an ideally straight line.

According to the square-bond model the axial ratio c/a depends upon the ratio of the ionic-bond length to the covalent-bond length. In Sec. IIIB, above, it was shown that the amount of covalence in the octahedral site, and therefore the bond length, depends on the type of bond in the tetrahedral interstices. If strong tetrahedral, covalent bonds are formed in the tetrahedral sites, the octahedral bond becomes more ionic and its length is increased so that the c/a ratio is decreased. The magnitude of this decrease is estimated in Appendix I.

The transition temperature at which the phase change from tetragonal to cubic symmetry occurs varies with the magnitude of the electronic-excitation energy. Since the excitation energy varies with the magnitude of the axial ratio c/a, the square-bond model leads to the prediction that the transition temperature varies monotonically as the c/a of the tetragonal phase provided all other factors are constant.

In summary, the square-bond model leads to the following predictions. (a) If cations (such as Cu^{II}, Mn^{III}, Pd^{II}) which form stable (dsp^2) bonds occupy octahedral sites in an oxide with the spinel structure, then the lattice is distorted to tetragonal symmetry with $c/a>1$. (b) At least 25 percent of the octahedral cations must form square covalent bonds if lattice distortion due to a long-range order of bond orientation is to occur. (c) The magnitude of the axial ratio c/a increases with the number of cations in octahedral sites which form square bonds. However, strong covalence in tetrahedral sites reduces the covalence in octahedral sites and hence the axial ratio. (d) There is a narrow temperature interval through which c/a decreases rapidly to 1 to give cubic symmetry. (e) Through the transition temperature interval there is a marked increase in electric conductivity σ and a marked decrease in the excitation energy. (f) The transition temperature depends upon the magnitude of c/a at absolute zero. (g) The transition shows a thermal hysteresis. (h) The excitation energy decreases with increasing temperature.

2. *Application of Theory to* $CuFe_2O_4$

Bertaut[27] has measured the lattice parameter and the cation-ordering parameter λ in $CuFe_2O_4$ as a function of temperature. The parameter λ is the fraction of octahedral interstices occupied by M^{2+} ions; the ferrospinel formula is therefore written as $M_{1-2\lambda}Fe_{2\lambda}(Fe_{2-2\lambda}M_{2\lambda})O_4$. The values $\lambda=0$, $\frac{1}{3}$, $\frac{1}{2}$ correspond, respectively, to a normal, statistically disordered, and inverse spinel. Bertaut observed that $CuFe_2O_4$ is tetragonal below 760°C. The axial ratio c/a and the parameter λ both decreased monotonically from room temperature to 700°C. Above 760°C the Cu^{2+} cations are randomly distributed in the anion interstices, and the lattice is cubic. It is especially interesting that Bertaut, who was convinced that the distortion to tetragonal symmetry was due to cation ordering, was unable to find any x-ray superstructure lines. The only ordering through the transition is a concentration of Cu^{2+} ions in the octahedral sites; among the octahedral sites, however, these ions are randomly distributed. Since the distortion is not accompanied by ion ordering among octahedral sites, it appears that the distortion is due to square-bond formation. Since the copper ions in CuO are Cu^{II} and form square, coplanar bonds, they undoubtedly form similar bonds in the spinel lattice. Because of the steric constraints, these bonds can only be formed if the Cu^{II} is in an octahedral site. If square covalent bonds are formed in an octahedral site, the elementary octahedron is distorted from cubic to tetragonal symmetry, and the alignment of these octahedra produces the macroscopic tetragonal symmetry of the crystal. This hypothesis is consistent with the observation of van Santen and van Wieringen[15] that in $CuFe_2O_4$ the oxygen octahedron surrounding the Cu^{2+} ion is tetragonally distorted.

[27] E. F. Bertaut, J. phys. radium **12**, 252 (1951).

The correlation between the axial ratio and the concentration of copper in octahedral sites as functions of temperature can be explained on the basis of the square-bond model. At room temperature Bertaut found 88 percent of the copper in octahedral sites ($\lambda=0.44$) and an axial ratio $c/a=1.06$. At 700°C the concentration of copper in octahedral sites had decreased to 74 percent ($\lambda=0.37$) and the axial ratio to $c/a=1.035$. The fact that Cu^{II} migrates from octahedral to tetrahedral sites indicates that Cu^{II} can be readily accommodated in tetrahedral sites although, because of square-bond formation, it is more stable in octahedral sites. The increase in entropy due to this migration apparently offsets the corresponding decrease in internal energy. That the decrease in the number of square bonds in octahedral sites due to the Cu^{II} migration is accompanied by a decrease in the distortion is consistent with the covalent model.

Quantitative correlation between Bertaut's observed axial ratios and λ produces the relation,

$$(\lambda-0.26)=3(c-a)/a \text{ for } \lambda\geq 0.26. \qquad \text{(IV.4)}$$

The factor $(\lambda-0.26)$ is a measure of the stress induced by square-bond formation. The strain is given by $(c-a)/a$. Equation (IV.4) expresses a linear stress-strain relationship (Hooke's law). It also states that in $CuFe_2O_4$ it is necessary to have 26 percent of the octahedral sites occupied by Cu^{II} before tetragonal distortion sets in. This corresponds closely to the estimate that at least 25 percent of the octahedral cations must form square bonds. Equation (IV.4) can be used to extrapolate to the case of a completely inverse $CuFe_2O_4$, for which $\lambda=0.5$. In this case half the octahedral sites contain Cu^{II}, and Eq. (IV.4) gives $c/a=1.08$. In a hypothetical spinel in which all octahedral sites are occupied by Cu^{II} ions and in which no covalent bonding exists in tetrahedral sites, there are twice as many square bonds in the octahedral sites so that $c/a=1.16$. This latter value is significant for comparison with the axial ratios of other tetragonal spinels which are thought to have square-bond formation in all the octahedral interstices.[28]

C. Tetragonal Spinels Containing Manganese

Hetaerolite ($ZnMn_2O_4$) and hausmannite (Mn_3O_4) each form a spinel lattice with tetragonal symmetry. Their axial ratios are reported as $c/a=9.23$ A/8.08 A $=1.14$,[29] and $c/a=9.42$ A/8.14 A $=1.16$,[29,30] respectively. Mason[29] and Romeijn[14] have determined that the Zn^{II} cations of hetaerolite are in the tetrahedral interstices. This cation arrangement corresponds to the optimum configuration for covalent bonding (see Table II) as well as for Madelung energies. Since the Mn^{II} cations may also form partially covalent bonds in tetrahedral interstices, it may be analogously assumed that in the hausmannite lattice Mn^{II} cations occupy tetrahedral interstices. This means that there are only two possible cation arrangements for these tetragonal spinels, *viz.*, $Zn^{2+}(Mn^{2+}Mn^{4+})O_4$ or $Zn^{2+}(Mn_2^{III})O_4$ for hetaerolite and $Mn^{2+}(Mn^{2+}Mn^{4+})O_4$ or $Mn^{2+}(Mn_2^{III})O_4$ for hausmannite. If the former arrangement is correct, then the distortion to tetragonality might be due to an ordering of Mn^{2+}, Mn^{4+} cations in alternating planes perpendicular to the c axis. Such a model was first proposed for hausmannite by Verwey.[13] If the latter arrangement is correct, then some mechanism other than ionic ordering, such as square-bond formation, must be responsible for distorting the lattice to tetragonal symmetry. Since it is impossible to distinguish between Mn^{2+}, Mn^{3+}, and Mn^{4+} by x-rays, it is necessary to use other measurements to distinguish between the two alternatives. Before the predictions on the basis of the two models are compared with various measurements, however, it should again be pointed out that Mn^{3+} has the electron configuration d^4 which, according to Sec. III, is particularly suited for the hybridization of empty (dsp^2) orbitals for coordinate-covalent square-bond formation. The square-bond model is, therefore, particularly satisfactory for spinels which contain Mn^{3+}, or rather Mn^{III}. Seven physical properties are discussed below.

1. Tetragonal Symmetry

The distortions to tetragonal symmetry in both hetaerolite and hausmannite must be caused by the same mechanism. Since Zn^{II}, which is well known to form tetrahedral covalent bonds, is tetrahedrally coordinated in hetaerolite, the tetrahedral-site bonds are certainly largely covalent. Therefore if the lattice distortion from cubic symmetry were due to an ordering of Mn^{2+} and Mn^{4+} ions in the octahedral interstices, the lattice would have orthorhombic symmetry as in Fe_3O_4. However, the lattice has tetragonal symmetry.

If the square-bond model is correct, then hetaerolite, which has strong covalent bonding in both the tetrahedral and octahedral interstices, should have a smaller axial ratio than hausmannite which has strong covalent bonds only in the octahedral sites. The calculation of Appendix I shows that a reduction from $c/a=1.16$ to $c/a=1.14$ is the order of axial-ratio reduction to be expected.

Further, an axial ratio $c/a=1.16$ for hausmannite, which forms square bonds in each occupied octahedral site and no covalent bonds in the tetrahedral sites, is exactly the ratio which is predicted from the extrapolation of the $CuFe_2O_4$ data.

2. Transition Temperatures

The transition of Mn_3O_4 from tetragonal to cubic symmetry was first observed at 1170°C by McMurdie

[28] McGuire *et al.*, Ceram. Age **60**, 22 (1952), report that $CuCr_2O_4$ is also an inverse, tetragonal spinel with $c/a=1.1$. Its axial ratio varies with temperature in an analogous manner to that of $CuFe_2O_4$.

[29] B. Mason, Am. Mineralogist **32**, 426 (1947).

[30] G. Aminoff, Z. Krist. **64**, 475 (1926).

and Golovato[31] who used a high-temperature x-ray technique. Romeijn[14] has measured the dc conductivity of Mn_3O_4 and $ZnMn_2O_4$ and found a conductivity discontinuity in the temperature intervals 1075°C $\leq T \leq$ 1150°C and 950°C $\leq T \leq$ 1025°C for the respective materials.

Although the Madelung bonding energy of the completely ordered state of Fe_3O_4 is 39.6 kcal/mole[10] greater than that of the disordered state, whereas the thermal energy corresponding to a transition temperature of $T=120$°K is 0.23 kcal/mole, the retention of considerable short-range order above 120°K can account for this discrepancy. If the phase transition to cubic symmetry were due to the same mechanism in Mn_3O_4 as in Fe_3O_4, then the transition temperature should be approximately four times higher in the 2–4 spinels since the Madelung-energy change due to octahedral-site ordering is four times as great. There is no reason to assume that the retention of short-range order above the transition temperature should be strikingly different in Mn_3O_4 than in Fe_3O_4. The analogous transition temperature for Mn_3O_4 would be 480°K, a value which is only a third of the observed transition temperature of 1443°K.

Since the covalent model offers an entirely different mechanism for lattice distortion, no correlation between the transition temperature of Fe_3O_4 and Mn_3O_4 is to be expected. There should, however, be a correlation of transition temperatures among the tetragonal spinels. On the basis of the square-bond model, the transition temperature varies with the magnitude of the axial ratio. This prediction is consistent with observations on $CuFe_2O_4$, $ZnMn_2O_4$, and Mn_3O_4 whose transition temperatures are 760°C, 1025°C, and 1170°C, corresponding to axial ratios 1.06, 1.14, and 1.16, respectively.

3. *Conductivity Discontinuity*

Romeijn[14] has measured the conductivity of Mn_3O_4 and $ZnMn_2O_4$ as a function of temperature and found that there is a conductivity and excitation-energy discontinuity associated with the phase change from tetragonal to cubic symmetry. Although there is a similar effect in Fe_3O_4, the discontinuity in magnetite occurs over a 1°C interval and exhibits no thermal hysteresis, whereas the manganate discontinuities occur over a 75°C interval, are considerably less pronounced, and exhibit a marked thermal hysteresis.

This contrast again corroborates the hypothesis that the phase transition in magnetite is due to a different mechanism than that in the manganates. The observed thermal hysteresis and resistivity discontinuity are greater in Mn_3O_4 than in $ZnMn_2O_4$. Also, the excitation-energy discontinuity is greater in Mn_3O_4. Such a correlation was predicted in Sec. B by means of the covalent model.

[31] H. F. McMurdie and E. Golovato, J. Research Natl. Bur. Standards **41**, 589 (1948).

4. *Resistivity at Room Temperature*

Verwey and de Boer[4] investigated the electrical resistivity of hausmannite at room temperature and found that it was 10^8 to 10^9 that of Fe_3O_4. Romeijn reported $\rho(Mn_3O_4)/\rho(Fe_3O_4) \approx 10^7$ ohm-cm/5×10^{-3} ohm-cm $=2\times10^9$. The relatively high conductivity of $Fe^{3+}(Fe^{2+}Fe^{3+})O_4$ is due to the transfer of electrons from octahedral Fe^{2+} to octahedral Fe^{3+} ions. In $Mn^{2+}(Mn_2^{III})O_4$ there is no exchange of electrons possible, and the observed high resistivity follows.

5. *Paramagnetism of* Mn_3O_4

If Mn_3O_4 consists of $Mn^{2+}(Mn^{2+}Mn^{4+})O_4$, there should be an antiferromagnetic coupling between the octahedral and tetrahedral Mn^{2+} cations which is comparable in strength to that between the Fe^{3+} cations in Fe_3O_4. However, Mn_3O_4 is paramagnetic at room temperature, and Millar[32] has measured the specific heat from 72°K to room temperature and found no λ-point in this range.

It will be shown in Sec. V, below, that paramagnetism in Mn_3O_4 follows from the covalent model.

6. γ-Mn_2O_3

The oxide γ-Mn_2O_3 is a spinel with 8/3 cation vacancies per unit cell. It has tetragonal symmetry with an axial ratio $c/a=1.16$.[4] This is the same axial ratio as that found in Mn_3O_4; the mechanism which produces the distortion to tetragonal symmetry is apparently the same in each case.

If the distortion mechanism consists of an ordering of Mn^{2+} and Mn^{4+} ions in the octahedral interstices, then γ-Mn_2O_3 must have a unit cell with eight Mn^{4+} ions in octahedral sites and (8/3) Mn^{4+} ions in tetrahedral sites. The (32/3) Mn^{2+} would be distributed over the remaining interstices of each type. Such a configuration, however, would minimize neither the Madelung energy, which favors all the Mn^{4+} ions in octahedral sites, nor any covalent-energy contribution due to tetrahedral-bond formation, which favors Mn^{2+} ions in the tetrahedral interstices. It is difficult to see, therefore, how such a cation distribution could possibly represent the minimum-energy configuration.

In the covalent model, on the other hand, all the cations are Mn^{3+}, or Mn^{III}, and the distortion is due to square-bond formation in the octahedral sites. If the Mn^{III} form square, quadricovalent bonds in the octahedral sites, the manganese ions are more stable in these lattice positions, and the correct formula for γ-Mn_2O_3 is $Mn_{2/3}(Mn_2)O_4$. All the vacancies are in the tetrahedral sites, and the axial ratio should, as it does, equal the value 1.16 which was calculated in Sec. IV B. Finally, the observed high electric resistivity[4] must follow as the covalent bond is full, and there is only one type of cation present.

[32] R. W. Millar, J. Am. Chem. Soc. **50**, 1875 (1928).

7. *Lattice Parameters*

If the distortion to tetragonal symmetry in Mn_3O_4 is due to the formation of square, covalent bonds about Mn^{3+} ions in the octahedral interstices, then the a-axis is a measure of the covalent-bond length and the c-axis is a measure of the ionic-bond length. If the cubic symmetry above the transition temperature is due to a resonance of the coplanar bonds among the three biaxial planes of the octahedral interstices, then the lattice parameter above the transition temperature is the measure of a bond length which is $\frac{2}{3}$ covalent and $\frac{1}{3}$ ionic. At room temperature $c=9.42$ A and $a=8.14$ A, so that the equivalent cubic lattice parameter is $a_0=\frac{2}{3}a+\frac{1}{3}c=8.56$ A. The coefficient of thermal expansion is estimated as

$$\Delta a_0/a_0\Delta T\approx 10^{-5} \text{ per degree.}$$

Since $a_0=8.56$ A, this gives $\Delta a_0\approx 0.09$ A for $\Delta T=1100°C$. At the transition temperature, 1170°C, the lattice becomes cubic with a measured lattice parameter $a'=8.64$ A. The lattice parameter which is calculated from the assumptions of square-bond formation in $Mn^{2+}(Mn_2^{III})O_4$ is $a_0+\Delta a_0=8.65$ A. The agreement is satisfactory.

C. The Indates and Metallic Indium

Passerini[33] has measured the lattice parameters of $CaIn_2O_4$ and $CdIn_2O_4$. Each of these crystals was found to be tetragonal; their respective axial ratios are 1.12 and 1.14. The cation distribution in these indates has not yet been determined. $MgIn_2O_4$, on the other hand, was found by Barth and Posnjak[34] to be cubic and inverse.

It is known that Cd^{II} tends to occupy tetrahedral interstices where (sp^3)-covalent-bond formation is possible. It may therefore be assumed that $CdIn_2O_4$ is normal. Ca^{2+}, like Mg^{2+}, forms ionic bonds so that it might be expected to occupy octahedral sites as do the Mg^{2+} ions in $MgIn_2O_4$.

If the lattice distortions to tetragonal symmetry are due to cation ordering in the octahedral sites, then $MgIn_2O_4$, which is inverse and has the larger difference in ionic radii (ionic radii of Mg^{2+}, In^{3+}, Cd^{2+}, Ca^{2+} are 0.78 A, 0.92 A, 1.03 A, 0.99 A, respectively), should give a larger distortion than $CdIn_2O_4$ and $CaIn_2O_4$. On the contrary $MgIn_2O_4$ is cubic, and $CdIn_2O_4$ and $CaIn_2O_4$ are tetragonal.

On the other hand, since the axial ratio of $CdIn_2O_4$ is identical with that of $ZnMn_2O_4$, which probably has (sp^3) covalent bonding in the tetrahedral sites and (dsp^2) covalent bonding in the octahedral sites, it is pertinent to inquire whether $CdIn_2O_4$ can form the same type of bonds. Although gaseous In^{3+} has a full $3d$ shell so that square-bond formation appears improbable, a study of metallic indium suggests the opposite.

[33] L. Passerini, Gazz. chim. ital. **60**, 754 (1930).
[34] T. F. W. Barth and E. Posnjak, Z. Krist. **82**, 325 (1932).

Metallic indium is face-centered tetragonal with an axial ratio $c/a=1.076$.[35] Since the lattice is distorted from cubic symmetry, it is not possible to explain the distortion by the interaction of the conduction electrons with the lattice, i.e., by the interaction of the Fermi surface with the Brillouin-zone energy-discontinuity surfaces.[36] In γ-Mn, which is also face-centered tetragonal but with $c/a=0.937$,[37] it is possible to explain the distortion as due to magnetic-exchange interactions,[38] but such interactions are not present in indium. Also, the indium distortion is in an opposite sense from that in γ-Mn.

In the usual discussion of energy bands in metals the d, s, and p bands are considered separately. However, if the d, s, and p bands overlap, a hybrid band would be more stable. In metallic indium there is doubtless an overlap of d, s, and p bands. It is suggested, therefore, that there is a hybrid, covalent, (dsp^2) band in indium. This band is full with one electron per indium atom left over for metallic bonding (see Fig. 3). If there is a (dsp^2) band, it will be reflected in square covalent bonds about each indium atom. The four nearest neighbors of each indium atom are indeed located at the corners of a coplanar square. It is suggested, therefore, that just as in the case of tetragonal spinels, the distortion to tetragonal symmetry in metallic indium is due to the formation of square covalent bonds. Further evidence for this suggestion is given in Appendix III.

The study of metallic indium indicates that when indium is in a lattice, the $3d$ electrons can be active in bond formation. It is probably just as correct, therefore, to consider the outer-electron configuration of In^{3+} to be d^9s as to consider it d^{10}. When the steric constraints favor (dsp^2)-bond formation, the configuration d^9s makes one d wave function available for hybridization which can enter into coordinate covalence with four coplanar oxygens as illustrated in Fig. 4. If In^{III} does indeed form square bonds, then the square-bond model is applicable to the indates. In^{III} in an octahedral site tends to cause tetragonal distortion through covalent square-bond formation. Normal indates have, therefore, tetragonal symmetry due to covalent square-bond formation in the octahedral sites. On the other hand, In^{III} in a tetrahedral site tends to prevent tetragonal distortion because its strong (sp^3) bonds weaken the covalent bonds in octahedral sites. This was shown in Sec. IIIB, where it was stated that the oxygen p orbitals are directed toward that near-neighbor cation which forms

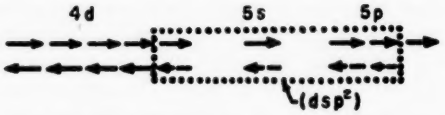


FIG. 3. Metallic indium with a full covalent (dsp^2) band.

[35] W. Betteridge, Proc. Phys. Soc. (London) **A50**, 519 (1938).
[36] J. B. Goodenough, Phys. Rev. **89**, 282 (1953).
[37] A. Westgren and G. Phragmen, Z. Physik **33**, 777 (1925).
[38] C. Zener, Phys. Rev. **81**, 440 (1951).

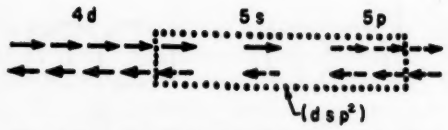


FIG. 4. Electron-spin configuration for In^{III} forming four coplanar covalent bonds.

the strongest covalent bond with oxygen. In inverse indates, tetragonal distortion is possible only if enough octahedral cations form sufficiently strong covalent square bonds.

Since $In(MgIn)O_4$ is cubic, there appears to be insufficient covalent bonding in octahedral sites for tetragonal distortion. There has to be a certain minimum number of covalently bound ions in octahedral sites for tetragonal distortion. This number depends on the strength of these covalent bonds. In $CuFe_2O_4$ the minimum number was shown [Eq. (IV.4)] to be 26 percent of the octahedral cations. Since in $In(MgIn)O_4$ the octahedral covalent bonds are weaker than in $CuFe_2O_4$, the minimum fraction of covalently bound ions in octahedral sites should here be greater than 26 percent. In $In(MgIn)O_4$ each oxygen is surrounded by one tetrahedral indium and three randomly distributed octahedral indium and magnesium ions. Magnesium is bound electrovalently. One oxygen can be bound covalently to one tetrahedral and two octahedral indium ions through the admixture of oxygen (sp^3) orbitals (Sec. IIIB, and Appendix I). Since only 50 percent of the oxygen ions have two octahedral indium nearneighbors, only 50 percent of the octahedral indium ions form significantly strong square bonds. This means that only 25 percent of the octahedral cations can form square bonds; as shown above this is insufficient to cause tetragonal distortion.

It is concluded that all normal indates are tetragonal, and that the inverse indates are cubic if the bivalent cation is electrovalent. If the bivalent cation can form covalent square bonds, the indate should be tetragonal even when inverse. Therefore it is predicted that if $CuIn_2O_4$ or $PdIn_2O_4$ form spinels, they will be found to be tetragonal, whether normal or inverse.

If the square-bond model is accepted for indates, the axial ratio of $CdIn_2O_4$ must equal that of $ZnMn_2O_4$, as has been observed. In each case the spinel is normal, and covalent bonds are formed in both tetrahedral and octahedral sites, so that the axial ratio $c/a=1.14$ is less than that of Mn_3O_4 (as explained in Appendix I).

Finally, the axial ratio of $CaIn_2O_4$, $c/a=1.12$, is greater than the extrapolated value for completely inverse $CuFe_2O_4$, $c/a=1.08$. Therefore this spinel cannot be inverse. This conclusion finds further support in the fact that if it were inverse, it would not be distorted because Ca^{2+} forms electrovalent bonds. The axial ratio $c/a<1.16$ is due to incomplete covalent bonding in the octahedral sites. If this indate were completely normal, this weakening could not be due to covalent bonding in tetrahedral sites. It is concluded, therefore, that $CaIn_2O_4$ is partially inverse; the concentration of covalently bound cations in octahedral sites is then less than that in the spinels with axial ratios 1.16 and 1.14, and the presence of In in both octahedral and tetrahedral sites causes a further reduction in axial ratio.

It is significant that gallates are cubic and that metallic gallium does not have the close-packed, tetragonally distorted structure. Apparently gallium, like mercury, tends to form pairs of atoms, presumably through $(1/3)s+(1/\sqrt{3})p_z+(\sqrt{5}/3)d_z$ hybrid orbitals.

H. Summary

In Table III there is a summary of the properties of those spinels which become orthorhombic or tetragonal. The orthorhombic-Fe_3O_4 transition is seen to be quite separate in character from the tetragonal transitions. The orthorhombic distortion has been definitely established as due to ordering of the octahedral Fe^{2+} and Fe^{3+} cations. The distortions due to ionic ordering are believed to be orthorhombic rather than tetragonal.

TABLE III. A comparison of some physical properties of spinels which distort from cubic symmetry below a transition temperature T_t. σ_a/σ_b and $(\epsilon_a-\epsilon_b)$ are the ratio of conductivity and the difference in electron activation energy above and below T_t.

Composition	Spinel distortion	Axial ratio	T_t	σ_a/σ_b	$\epsilon_a-\epsilon_b$ in ev	Thermal hysteresis
Fe_3O_4	Orthorhombic	$a:b:c=$ 5.912:5.945:8.388[a]	−153°C[b]	90[b]	−0.05[b]	None[b]
Mn_3O_4	Tetragonal	$c/a=1.16$[c,d]	1170°C[c]	5[c]	−0.55[c]	~45°C[c]
$CuFeO_{2+\delta}$	Tetragonal	$c/a=1.16$[f,g]	760°C[g]	...	...	...
$CuCr_2O_4$	Tetragonal	$c/a=1.1$[h]	...	...	...	...
γ-Mn_2O_3	Tetragonal	$c/a=1.16$[i]	...	...	...	...
$ZnMn_2O_4$	Tetragonal	$c/a=1.14$[d]	1025°C[e]	2[e]	−0.4[e]	~40°C[e]
$CdIn_2O_4$	Tetragonal	$c/a=1.14$[j]	...	...	...	...
$CaIn_2O_4$	Tetragonal	$c/a=1.12$[j]	...	...	...	...

[a] S. C. Abrahams and B. A. Calhoun, Acta Cryst. **6**, 105 (1953).
[b] B. A. Calhoun, Phys. Rev. **94**, 1577 (1954).
[c] McMurdie, Sullivan, and Mauer, J. Research Natl. Bur. Standards **45**, 35 (1950).
[d] B. Mason, Am. Mineralogist **32**, 426 (1947).
[e] F. C. Romeijn, Philips Research Repts. **8**, 304 (1953).
[f] Extrapolated value for every octahedral site occupied by a Cu^{2+} cation. Measured values are $c/a=1.06$ and 1.035 at 20°C and 700°C, respectively.
[g] E. F. Bertaut, J. phys. radium **12**, 252 (1951).
[h] McGuire, Howard, and Smart, Ceram. Age **60**, 22 (1952).
[i] E. J. W. Verwey and J. H. de Boer, Rec. trav. chim. **55**, 531 (1936).
[j] L. Passerini, Gazz. chim. ital. **60**, 754 (1930).

This necessarily follows if there is covalent bonding in the tetrahedral sites which resists a distortion of its coordination angles.

The tetragonal symmetry in Mn_3O_4, γ-Mn_2O_3, $ZnMn_2O_4$, $CuFe_2O_4$, $CuCr_2O_4$, $CdIn_2O_4$, and $CaIn_2O_4$ is believed due to the formation of an ordered system of four coplanar, covalent bonds and two linear, ionic bonds in the octahedral sites which are occupied by Mn^{III}, Cu^{II}, or In^{III} cations. If the indate is inverse, the indium ions which are in the tetrahedral sites bond covalently; those in the octahedral sites form an ordered array of square bonds only if the bivalent ions also form covalent bonds. The lattice is therefore cubic if the bivalent cation is electrovalent. The tetragonal symmetry of metallic indium is also believed due to the formation of a (dsp^2) covalent bond.

Finally it is believed that there are no Mn^{4+} cations in Mn_3O_4, γ-Mn_2O_3, and $ZnMn_2O_4$.

V. MAGNETIC EXCHANGE

A. Exchange Mechanisms

Three mechanisms, known as direct exchange, double exchange, and superexchange, have been used to explain the interaction between electron spins. These interactions are based on the Pauli exclusion principle. Direct exchange tends to align the net spin on neighboring lattice elements antiparallel to each other. In spinels cations are separated by anions without net spin so that direct exchange is negligible.

Double exchange, as proposed by Zener,[39] requires the transfer of electrons between two kinds of ions, particularly two of the same element but with different charges, and hence is inhibited if these cations are ordered. Double exchange tends to align spins parallel to each other; also it produces a relatively high conductivity.

Kramers[40] suggested the third mechanism, called superexchange. Superexchange is due to the admixture of excited states with the ground state of a system containing two like cations separated by an anion. The excited states are the result of removing an electron from the oxygen ion O^{2-} and placing it in an empty or half-filled cation orbital. If the cation orbital is empty, the transferred electron aligns its spin parallel to the net spin already on the cation in accordance with Hund's rule. If the orbital is already half filled, the transferred electron aligns its spin antiparallel to that of the electron already there, and hence to the existing net spin on the cation. The O^{1-} ion from which the electron is transferred has a net spin which is antiparallel to that of the removed electron. The net spin on the oxygen ion is therefore antiparallel to the spin of the reduced cation if the cation d shell was originally less than half filled; parallel, if the cation d shell was originally at least half filled. These alignments are shown in Fig. 5.

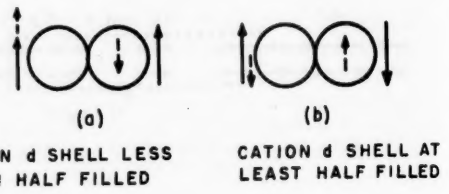


FIG. 5. Spin alignment in superexchange.

According to the superexchange model the net spin on the O^{1-} ion interacts directly with that of a neighboring cation on the other side. It was predicted, therefore, that the two cations orient their spin parallel if their d shells are less than half filled, antiparallel if the d shells are more than half filled.

Kramers' model is inconsistent in two ways with the concepts of covalence in spinels which are proposed in this paper. In the first place it was assumed that when the cation d shell is at least half filled, any additional electron will occupy a half-filled d orbital and align its spin antiparallel to that existing on the cation. The Pauli principle would indeed appear to predict such an alignment. It should be recalled, however, that this prediction is based on the supposition that there are no unfilled orbitals available that are degenerate with the $3d$ orbitals. When a new electron is added to a cation, it can either occupy an empty orbital or a half-filled one. In the former case energy may be required to raise the electron to the energy of the empty orbital, but this energy may be offset when the new electron aligns its spin parallel to the existing spin on the cation. In the latter case no energy is required to raise the electron to a higher level, but no exchange energy is gained, because the electron now must align its spin antiparallel to the existing spin. Therefore the spin alignment in superexchange depends on whether the exchange energy is larger or smaller than the difference in energy between d orbitals and the next level. It has been shown that, because of covalent overlap, there are indeed hybrid orbitals whose energy is about the same as that of the filled, or partially filled, d orbitals. Therefore superexchange would always produce parallel alignment when hybrid cation orbitals are occupied because, in effect, the d shell is extended by s and p orbitals of the same energy and is therefore less than half filled.

In the second place Kramers assumed the transfer of an electron from an oxygen ion to a cation. Somehow this transferred electron must retain its identity as an oxygen p electron since it is assumed that the spin on the resultant O^{1-} ion remains antiparallel to that of the transferred electron. Without covalence it is not at all clear why an electron should be simultaneously identified with both a cation and an anion. On the other hand, if the excited state of the system is covalent, this would indeed be the case.

In crystals both sides of an O^{2-} p orbital may overlap with an empty cation orbital. For coordinate covalence

[39] C. Zener, Phys. Rev. 82, 403 (1951).
[40] H. A. Kramers, Physica 1, 182 (1934).

the overlapping O^{2-} orbital and empty cation orbital should be nearly degenerate. In ionic crystals this is not the case for the ground state of the ion. However, if the cation in an electrovalent crystal has a net spin, exchange energy is gained if one of the two electrons in the oxygen p orbitals is excited to become part of the cation. Since only one electron participates in this bond, the bond is called "semicovalent." "Semicovalence" is thus defined as a bond due to the coupling of a single anion electron to the net spin of a cation. Fe^{3+} and Mn^{2+}, which have been observed to favor tetrahedral sites in the spinel lattice, are thought to be stabilized in these sites through semicovalence. In Sec. IIIC, 1, the term "partially covalent" was used to explain this stabilization.

Obviously the electron that participates in semicovalence has a spin parallel to that of the cation. The other anion electron therefore has a spin antiparallel to the spin of that cation. This other electron may participate in semicovalence with a cation on the other side of the anion. In that case the spins of the two cations would be coupled antiparallel to each other. This model provides a mechanism for antiferromagnetic coupling between two cations separated by an anion. This mechanism is called "semicovalent exchange." Its strength depends on the overlap between the O^{2-} p orbitals and the cation orbitals, and therefore on the cation-oxygen-cation angle. If semicovalence occurs with one cation only, the interaction between the anion and other cations is confined to direct exchange, and weak ferromagnetic coupling may result.

In summary, since the superexchange model has only been tested in the case of ferromagnetism and antiferromagnetism (ferromagnetic coupling can be explained by other mechanisms, such as double exchange), it is suggested that semicovalent exchange is a more consistent model for the antiferromagnetic-exchange interactions which have previously been attributed to superexchange. There are three conditions which favor semicovalent exchange: the formation of predominantly electrovalent bonds, cations with empty hybrid orbitals which overlap with a full anion p orbital, and a location of interacting cations at opposite sides of an anion.

B. Magnetic Properties of Some Oxides

There are two exchange mechanisms which are important in predominantly electrovalent metallic oxides, double exchange and semicovalent exchange. Double exchange, which causes ferromagnetic alignment and increased conductivity, occurs between cations of like atoms, but different ionization, if they are randomly distributed on equivalent lattice sites. In a spinel lattice there is negligible double-exchange interaction between tetrahedral- and octahedral-site cations.

Semicovalent exchange couples cation moments antiferromagnetically. This coupling is strongest when the cation-anion-cation bond is 180°. In the spinel lattice neighboring octahedral cations make a 90° angle with their common anion, and the semicovalent exchange between them is small compared to that between the tetrahedral and octahedral cations. These mechanisms are applied below to several metal oxides, and in each the semicovalent-exchange mechanism is compared with superexchange.

It should be noted that cations which form coordinate-covalent bonds do not participate in either semicovalent exchange or double exchange. Such ions are therefore paramagnetic.

1. Some Antiferromagnetic Crystals

In oxides which have cations in one valence state only, double exchange cannot occur. If the metal is a transition element, semicovalent exchange is possible. Many such lattices are known to be antiferromagnetic. A typical example is MnO which has the rock-salt structure. Whereas superexchange would require the formation of $Mn^{1+}O^{1-}$ in an excited state, semicovalent exchange couples the system antiferromagnetically through electron pairs whose orbitals contain an admixture of excited $Mn^{II}O^{II}$ states.

It is interesting that the Curie points of manganous compounds with the same rock-salt structure increase with the substitution of anions of lower electronegativity, *viz.*, MnO (122°K), MnS (165°K), MnSe (247°K), MnTe (307°K).[41] Such a relation is consistent with a model which predicts that the strength of the antiferromagnetic coupling is increased as the admixed excited states become more nearly degenerate with the electrovalent ground state. It is consistent, therefore, with both superexchange and semicovalent exchange.

If one compares the Curie points of the oxides of Mn, Fe, Co, and Ni, all of which have the rock-salt structure, a similar progression of Curie temperatures is noted, *viz.*, MnO (122°K), FeO (198°K), CoO (291°K), NiO (523°K). On the basis of a superexchange model this progression can be attributed to the progressive increase in the ionization potential of the bivalent ion. On the basis of a semicovalent model it may be attributed to a progressive decrease in the energy of the hybrid (d^2sp^3) orbitals. The high Curie point of NiO indicates that semicovalence has a significant stabilizing effect on Ni^{2+} in octahedral sites of O^{2-} ions. Although there is no normal covalence in the Ni^{2+}—O^{2-} bond, semicovalence is possible. The apparent preferance of Ni^{2+} for octahedral sites in the spinel lattice[14] may indeed be due to this stabilization. Further, this stabilization may well explain the normal spinel structure of $Ge(Ni_2)O_4$ and $Ge(Co_2)O_4$ as contrasted with the olivine structure of $Ge(Mg_2)O_4$. Since Mg^{2+}, Co^{2+}, and Ni^{2+} have comparable ionic radii, this difference in crystal structure cannot be understood from any previous model.

[41] H. Bizette, J. phys. radium **12**, 161 (1951).

2. Fe_3O_4

Fe_3O_4 is an inverse spinel, $Fe^{3+}(Fe^{2+}Fe^{3+})O_4$. The interaction between the tetrahedral Fe^{3+} ions and the octahedral ions can be fully explained by semicovalent exchange but only partially by superexchange. According to both models the tetrahedral Fe^{3+} ions would couple antiferromagnetically with the octahedral Fe^{3+} ions. However, superexchange does not provide for the $Fe^{3+}-Fe^{2+}$ interaction since it applies only to interactions between identical ions. Semicovalent exchange corresponds to antiferromagnetic coupling in this case also. Between the low-temperature transition and the Curie point the spins on the octahedral sites are also aligned parallel to each other through double exchange. Therefore all the spins of the octahedral-site cations are aligned parallel to one another and antiparallel to those of the tetrahedral-site cations. In this temperature interval both the superexchange and semicovalent-exchange mechanisms provide for the ferromagnetism of Fe_3O_4. Below 119°K, however, the Fe^{3+} and Fe^{2+} cations are ordered so that double exchange cannot occur. At these low temperatures the superexchange model provides no mechanism for a parallel alignment of all the octahedral-site spins. Since Fe_3O_4 is ferromagnetic below 119°K, the antiparallel alignment of all octahedral-site spins with respect to the tetrahedral-site spins, and the consequent parallel alignment of the octahedral-site spins must be due to semicovalent exchange. Semicovalent exchange provides the more adequate model.

3. Mn_3O_4

It is proposed in this paper that the structure of Mn_3O_4 is $Mn^{2+}(Mn_2^{III})O_4$. Elsewhere the structure $Mn^{2+}(Mn^{2+}Mn^{4+})O_4$ has been proposed. Mn^{2+} in tetrahedral sites is partially covalent. Mn^{III} forms square covalent bonds. Octahedral Mn^{2+} and Mn^{4+} would be ionically bound. In the $Mn^{2+}(Mn^{2+}Mn^{4+})O_4$ model Mn^{2+} octahedral-tetrahedral interactions would be antiferromagnetic according to both superexchange and semicovalent exchange. The (Mn^{2+} tetrahedral)-(Mn^{4+} octahedral) interaction cannot be treated with superexchange, but semicovalent exchange predicts antiferromagnetic exchange. Thus the structure $Mn^{2+}(Mn^{2+}Mn^{4+})O_4$ would be ferromagnetic just like $Fe^{2+}(Fe^{2+}Fe^{3+})O_4$. Yet down to 72°K Mn_3O_4 has been found to be paramagnetic.

In the structure $Mn^{2+}(Mn^{III})O_4$ there would be no octahedral-tetrahedral semicovalent exchange because all octahedral-site cations are covalently bound. There would be no double exchange because Mn^{2+} and Mn^{III} are ordered in different types of sites. Thus the model $Mn^{2+}(Mn_2^{III})O_4$ is paramagnetic. Since Mn_3O_4 has been found to be paramagnetic as low as 72K°, this observation supports the $Mn^{2+}(Mn_2^{III})O_4$ model in preference to the $Mn^{2+}(Mn^{2+}Mn^{4+})O_4$ model.

4. $CuFe_2O_4$

This spinel is partially inverse with the formula $Fe_2Cu_{1-2\lambda}(Cu_{2\lambda}Fe_{2-2\lambda})O_4$ as discussed in Sec. IV B, 2. If Cu^{II} forms square covalent bonds in the octahedral sites, as is suggested in this paper to explain the tetragonal structure of $CuFe_2O_4$, then the octahedral copper does not contribute to the magnetism.

Since the exchange between the octahedral- and tetrahedral-site cations is the only spin interaction of significance in $CuFe_2O_4$, it is expected that the octahedral-site spins, if they contribute at all to the magnetic moment, are parallel to one another and antiparallel to the tetrahedral-site spins. If $M_a\mu_B$ and $M_b\mu_B$ are the atomic moments of tetrahedral-site and octahedral-site copper respectively, where μ_B is a Bohr magneton, then the magnetic moment per molecule of $CuFe_2O_4$ is $|(2-2\lambda)5+2\lambda M_b-(2\lambda)5-(1-2\lambda)M_a|\mu_B = |(1-2\lambda)10-M_a+2\lambda(M_b+M_a)|\mu_B$. Pauthenet and Bochirol[42] report $2.19\mu_B$ for the magnetic moment of a $CuFe_2O_4$ sample which had been quenched from a 700°C anneal. Bertaut[27] found that a similarly treated $CuFe_2O_4$ sample had an ordering parameter $\lambda=0.37$. Although a direct comparison between data from two different, though similarly treated, samples is not strictly justified, it is interesting that for $\lambda=0.37$ the predicted magnetic moment of $CuFe_2O_4$ is $(2.6+0.74M_b-0.26M_a)\mu_B$ which, for $M_a=1$ and $M_b=0$, is $2.3\mu_B(>2.19\mu_B)$. It is reasonable to assume, therefore, that the octahedral-site copper ions do not contribute to the magnetic moment.

VI. SUMMARY

The covalent bond in spinels has been shown to influence cation ordering, lattice distortions, and magnetic-exchange interactions. Covalent bonding modifies the Madelung-energy and geometrical calculations of previous workers. Octahedral and square covalent-bond formation are as significant as is tetrahedral-bond formation in determining cation ordering in spinels. From only qualitative considerations it was possible to predict nearly all the observed cation distributions.

While orthorhombic distortion, as observed in Fe_3O_4, can be ascribed to cation ordering in octahedral sites, tetragonal distortion is shown to be due to ordering of the square bonds in octahedral sites. The two distortion mechanisms are vastly different, as is evident from the difference in transition temperature, thermal hysteresis, etc. A certain minimum concentration of square-bond-forming cations in octahedral sites must be exceeded for the lattice to distort. Covalent bonds in tetrahedral sites tend to hinder tetragonal distortion.

The structure of Mn_3O_4 was found to be $Mn^{2+}\times(Mn_2^{III})O_4$ rather than $Mn^{2+}(Mn^{2+}Mn^{4+})O_4$. The tetragonal structures of metallic indium and the indates are accounted for by the square-bond model.

[42] R. Pauthenet and L. Bochirol, J. phys. radium **12**, 249 (1951).

Semicovalence is defined as covalent bonding due to the coupling of a single electron spin to the net spin on a cation. This coupling leads to semicovalent exchange, a mechanism producing antiferromagnetic coupling of cation spins. The paramagnetism of $Mn^{2+}(Mn_2^{III})O_4$ is explained on the basis of semicovalent exchange, as is the magnetization of Fe_3O_4 below 119°K.

APPENDIX I. THE INFLUENCE OF OXYGEN (sp^3)-ORBITAL HYBRIDIZATION ON THE AXIAL RATIO OF TETRAGONAL SPINELS

In Sec. IIIB, it has been pointed out that the oxygen p^3 orbitals in a spinel are instrumental in the formation of covalent bonds between oxygen and its near neighbors in octahedral sites. Where there are empty cation orbitals in tetrahedral sites the oxygen can be bound covalently to at least two octahedral and one tetrahedral cation through the admixture of oxygen (sp^3) orbitals with the p^3 orbitals. The p^3 orbitals overlap strongly with the empty orbitals in the octahedral sites but not with those in tetrahedral sites while the set of (sp^3) orbitals that overlap strongly with a tetrahedral site do not overlap as strongly with the octahedral sites as do the p^3 orbitals. Therefore the admixture of oxygen (sp^3) wave functions with the p^3 wave functions increases the amount of covalent bonding of oxygen with a tetrahedral near neighbor but decreases the amount of covalent bonding with the octahedral near neighbors.

It has been proposed that the tetragonal distortion of many spinels is due to square covalent bonds in octahedral sites. When covalent bonds are formed by one oxygen with both octahedral and tetrahedral sites through (sp^3) admixture, the bonds are less covalent in the octahedral sites than they would have been if no covalent bonds with the tetrahedral ions existed. Therefore the axial ratio of a tetragonal spinel is a maximum when only the octahedral cations are covalently bound, and decreases with increasing amount of covalence in tetrahedral sites.

This can be quantitatively shown as follows. The amount of covalence in the octahedral sites due to oxygen p^3-orbital overlap and that in tetrahedral sites due to oxygen (sp^3)-orbital overlap is arbitrarily set at unity. The amount of covalence in the octahedral sites due to oxygen (sp^3)-orbital overlap is given by $\delta(<1)$, that in the tetrahedral sites due to p^3 overlap by the very small fraction ϵ. The wave function for the (sp^3) orbitals and the p^3 orbitals are admixed in the ratio $\alpha:\beta$, where $\alpha^2+\beta^2=1$. The normalized mixed wave function is then denoted by $\alpha(sp^3)+\beta(p^3)$. The amounts of covalence due to each term in the wave function is shown in Table IV. [The physical properties are proportional to the square of the wave functions; cross products $\alpha\beta(sp^3)(p^3)$ are neglected here.] Therefore the amount of covalence in octahedral sites is $(\alpha^2\delta+\beta^2)$. The amount of electrovalence is $1-(\alpha^2\delta+\beta^2)$. If it is assumed that the bond length is a linear function of the amounts of covalence and electrovalence, the following bond length is then derived:

$$a'=(\alpha^2\delta+\beta^2)a+(1-\alpha^2\delta-\beta^2)c, \qquad (A.1)$$

where a is the covalent bond length in octahedral sites in the absence of (sp^3) admixture, and c is the electrovalent bond length in octahedral sites. The axial ratio is then given by

$$\frac{c}{a'}=\frac{c}{(\alpha^2\delta+\beta^2)a+(1-\alpha^2\delta-\beta^2)c}. \qquad (A.2)$$

In spinels in which there is no appreciable covalent bonding in tetrahedral sites, such as Mn_3O_4, the axial ratio has been observed to be $c/a=1.16$. The amount of overlap of oxygen (sp^3) orbitals with octahedral cation orbitals is estimated to be about 70 percent of that of oxygen p^3 orbitals with octahedral cation orbitals. On the assumption that $\delta=0.8$, the axial ratio is plotted in Fig. 6 as a function of the admixture of oxygen (sp^3) orbitals.

For spinels like $CdIn_2O_4$ and $ZnMn_2O_4$ an axial ratio 1.14 has been observed; for Fig. 6 this is seen to correspond to $\alpha=0.6$, $\beta=0.8$. It is significant that such a large amount of admixture of (sp^3) orbitals, and hence considerable covalence in tetrahedral sites, only slightly reduces the axial ratio. In fact pure (sp^3) wave functions for oxygen ($a=1$, $\beta=0$) still allow enough covalence in the octahedral sites for distortion to an axial ratio $c/a=1.10$.

TABLE IV. Contributions of oxygen (sp^3) and p^3 orbitals to the covalence in tetrahedral and octahedral sites.

	Tetrahedral sites	Octahedral sites
$\alpha(sp^3)$	α^2	$\alpha^2\delta$
$\beta(p^3)$	$\beta^2\epsilon$	β^2

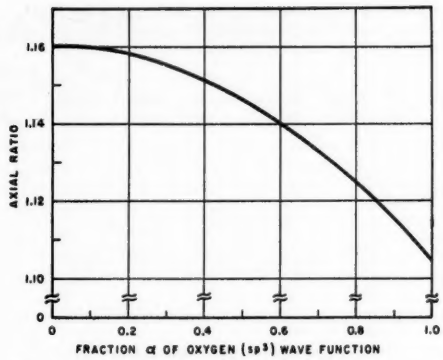


FIG. 6. The axial ratio as a function of the amount of admixture of oxygen (sp^3) orbitals.

APPENDIX II. CALCULATION OF THE TETRAGONAL DISTORTION IN Fe_3O_4 FROM AN IONIC MODEL

To calculate θ_1 and θ_1' of Eq. (IV.1), it is first necessary to obtain θ_2 and θ_2' which, from the law of cosines,

TABLE V. Self-consistent calculation of λ for Fe_3O_4.

λ	θ_1	θ_1'	ϵ	ϵ'	δ	δ'	Δ
1.020	1°57′	0°20′	0.0469	0.0095	0.0247	0.004	0.0013
1.024	2°10′	0°0′	0.0468	0.0095	0.0273	0	−0.0019

are given by

$$\cos\theta_2=\frac{\lambda^2+(1+\zeta)^2-(1-\zeta')^2}{\{(2\lambda^2+1)[2(1+\zeta)^2-1]\}^{\frac{1}{2}}},$$

$$\cos\theta_2'=\frac{\lambda^2+(1-\zeta')^2-(1+\zeta)^2}{\{(2\lambda^2+1)[2(1-\zeta'^2)-1]\}^{\frac{1}{2}}}.$$

The angles θ_1 and θ_1' then follow directly from the geometrical relation

$$\cos(\theta_1+\theta_2)=(2\lambda^2+1)^{-\frac{1}{2}}=\cos(\theta_2'-\theta_1').$$

If $(1+\zeta)^2=1.0486$, $(1-\zeta')^2=0.9905$, $d^2=0.75$, then the values listed in Table V follow where $\Delta=[d^2-\frac{1}{2}\times(1-\epsilon)^2]^{\frac{1}{2}}+[d^2-\frac{1}{2}(1+\epsilon')^2]^{\frac{1}{2}}-\lambda(1+\delta-\delta')$. The correct value of λ is that for which $\Delta=0$, or $\lambda=1.022$.

APPENDIX III. TETRAGONAL DISTORTION IN Cd-In ALLOYS

Betteridge[35] added small amounts of Cd to In and observed the change in lattice parameter. At 100°C, 14.5 atomic percent of Cd was soluble in In. The axial ratio decreased from $c/a=1.068$ for pure In to $c/a=1.0$ for 4.5 atomic percent Cd. At 20°C the solubility limit of Cd was 4.5 atomic percent. Between pure In and the solubility limit c/a decreased from 1.076 to 1.04.

Since the Cd atoms do not contribute to the square covalent bonds, the addition of Cd to In adds vacancies to the (dsp^2) band. The tetragonal distortion, and therefore the energy gap between the (dsp^2) band and the conduction band, is reduced. At 100°C and 4.5 atomic percent Cd enough covalent-band electrons have sufficient energy to go into the conduction band so that the covalent ordering is destroyed, and the lattice becomes cubic. At 20°C there are fewer covalent-band electrons which are excited into the conduction band, and the c/a ratio is higher, at all cadmium concentrations, than it is at 100°C. At 4.5 percent Cd and 20°C the covalent bonds are still strong enough to force the Cd out of solution. The tetragonal symmetry is therefore preserved.

If the phase transition from tetragonal to cubic symmetry occurs in 4.5 atomic percent Cd at 100°C, a similar transition might be expected a few degrees higher in pure In. The melting point of In is 156.4°C.[43] Apparently the low melting point of In occurs below the phase transition from tetragonal to cubic symmetry.

[43] Taylor Lyman, editor, *Metals Handbook, 1948 Edition* (American Society for Metals, Cleveland, 1948).

Etch Patterns and Ferroelectric Domains in $BaTiO_3$ Single Crystals

JOHN A. HOOTON AND WALTER J. MERZ
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received December 17, 1954)

It has been found that ferroelectric $BaTiO_3$ single crystals show a differential etch rate when etched in hydrochloric acid. The positive end of the electric polarization etches much faster than the negative one. Thus this phenomenon affords a valuable new tool for the investigation of the ferroelectric domains in $BaTiO_3$. Some interesting pictures of domain arrangements are shown and discussed. The results obtained demonstrate the extremely strong electric dipole-dipole coupling in the forward direction and the weak coupling in the directions perpendicular to it. This electrostatic interaction governs the domain patterns in ferroelectric $BaTiO_3$ crystals.

I. INTRODUCTION

IT has been found that $BaTiO_3$ single crystals show a differential etch rate which can be related to the ferroelectric domains. The most important result of this observation is that we can easily and quickly distinguish between the positive and the negative ends of the electric dipoles. We find that the positive end (head) of the dipole etches quite rapidly in hydrochloric acid while the negative end (tail) of the dipole etches very slowly, if at all. The usual optical methods have been very useful for determining the domain arrangements and the direction of the electric polarization. However, the sense of the polarization could only be determined by applying external electric fields or stresses to the crystal.[1] Thus we find that the optical methods and the etching technique complement each other perfectly. In the following paragraphs we will report some interesting results obtained by this new method.

II. PROCESSING OF SAMPLES

The method of preparing the crystals for study is extremely simple. The crystals are placed in a test tube and etched for one to ten minutes, depending upon the desired degree of development of the patterns, in concentrated hydrochloric acid at room temperature. The acid is then poured off the crystal rinsed in water and washed in ethyl alcohol. The microscopic patterns are studied using reflected light. Magnifications in the 100× to 200× range are most satisfactory.

III. ETCH PATTERNS OF POLED CRYSTALS

At room temperature $BaTiO_3$ shows tetragonal symmetry which is associated with a spontaneous polarization which may be thought of as a shift of the Ti ions with respect to the oxygen octahedra. The polarization direction lies parallel to the tetragonal *c*-axis. Regions of the crystal in which all the dipoles are aligned perpendicular to the crystal flat are called *c*-domains because we look along the polar *c*-axis. Similarly, regions in which all the dipoles are aligned in a direction parallel to the crystal surface are called *a*-domains. Optically it is very easy to distinguish between an *a*- and a *c*-domain. However, it has been a problem to determine the sense of a given polarization direction because parallel and antiparallel domains have the same optical properties and thus look the same. By applying external electric fields or stresses it was possible to determine the sense of the polarization.[1]

Etching experiments on dc poled crystals show that the etchant (HCl) attacks very rapidly at the positive end (head) of the polarization and quite slowly at the negative end (tail), as can be seen in Fig. 1. The islands which are antiparallel regions go right through the crystal. Figure 1 represents a poorly poled crystal since it consists of almost as many domains polarized in one direction as domains polarized in the opposite direction. A well-poled crystal usually shows just very few scattered circular areas. Whether these areas resist the poling voltage or whether they are regions which invert after the field has been removed and before the crystal has been etched, has not been determined. Figure 1 shows beyond any doubt the differential etch rate for the two ends of the electric dipoles.

The etch rate perpendicular to the dipole direction, that is for an *a*-domain, is intermediate between the rapid rate for the titanium end of the dipoles and the slow rate for the oxygen end of the dipoles. The different etch rates are shown in Fig. 2.

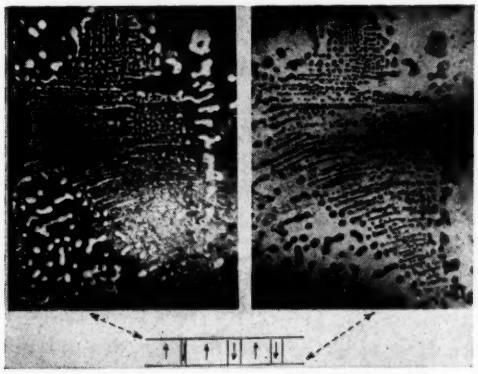


FIG. 1. Photograph of top and bottom of a poorly poled $BaTiO_3$ crystal showing many antiparallel polarized domains. Notice the perfect mirror image. (One print is reversed.)

[1] Walter J. Merz, Phys. Rev. 95, 690 (1954).

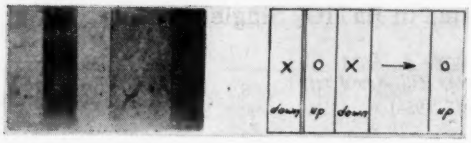


FIG. 2. Surface appearance of etch response of differently polarized regions.

IV. DOMAIN REFLECTIONS

Earlier investigations[1] have shown that the ferroelectric domains in $BaTiO_3$ crystals are always so arranged that there is no build-up of a space charge on the domain walls. This implies a very strong coupling between dipoles in the dipole direction whereas the coupling perpendicular to the dipole direction is weak. For example, we could show that the 90° walls between two *a*-domains are always along [011] directions and the subdomains in the two *a*-domains are arranged in such a way that the head of one dipole always joins the tail of the neighboring dipole as shown in Figs. 3 and 4.

We have etched some *c*-domain crystals containing a few *a*-domains which were so arranged as to illustrate clearly this strong linear coupling in the forward direction (head to tail coupling). This is shown in Fig. 5, where the dark areas are the regions where the positive ends of the dipoles are at the surface, the bright areas where the negative ends of the dipoles are at the surface and the grey areas are *a*-domains as explained before in Fig. 2. To the extreme left of Fig. 5 we see a distinctively shaped region (a) of inverted polarization. On the right side of the two very narrowly spaced sharp lines (two 90° walls) we see the mirror image (b) of this region. The inverted regions are oppositely poled with respect to each other and also to their background region. The two sharp lines represent the boundaries of the wedge-shaped *a*-domain through which the reflection of the oppositely polarized region takes place. Figure 5 thus shows in a striking way the strong coupling of the dipoles in the forward direction. The center part of Fig. 5 shows a second odd-shaped antiparallel polarized region (c) which is reflected through a wide *a*-domain to the right side of the picture (d). This reflection is the same type as the first one except that the path through the *a*-domain is large. The

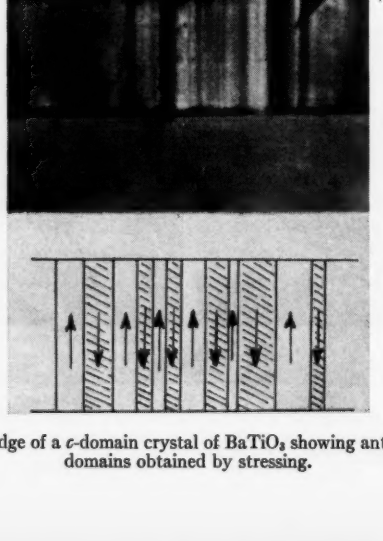


FIG. 3. Edge of a *c*-domain crystal of $BaTiO_3$ showing antiparallel domains obtained by stressing.

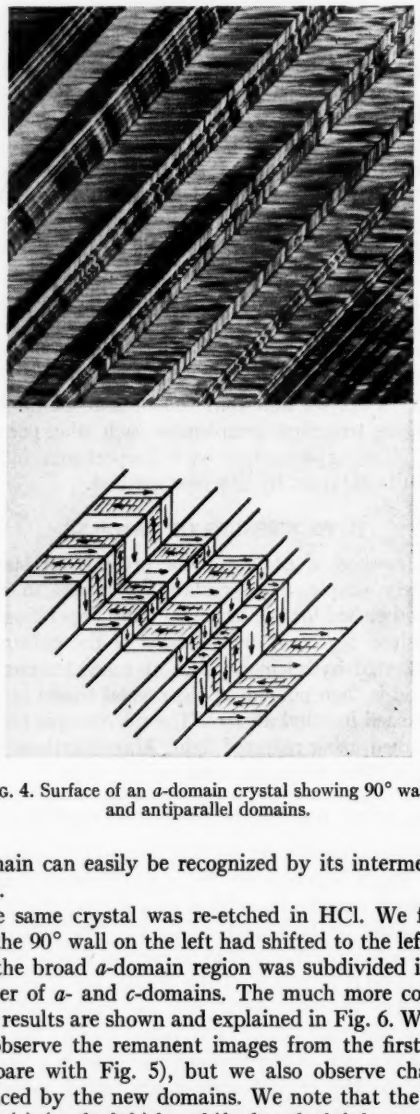


FIG. 4. Surface of an *a*-domain crystal showing 90° walls and antiparallel domains.

a-domain can easily be recognized by its intermediate shade.

The same crystal was re-etched in HCl. We found that the 90° wall on the left had shifted to the left and that the broad *a*-domain region was subdivided into a number of *a*- and *c*-domains. The much more complicated results are shown and explained in Fig. 6. We can still observe the remanent images from the first etch (compare with Fig. 5), but we also observe changes produced by the new domains. We note that the dark image (a) (to the left) has shifted to the left because the 90° wall has shifted. At the point where we observed two sharp lines before we observe now three lines. We

think this "triplet" is composed of the traces of the two original lines lying to the left and the new intersection of the same wall with the surface which has moved to the right down the 45° slope of the wall. This displacement to the right is equal to the amount of material etched away in the second etching. Thus the region between the single line to the extreme left and the third line of the triplet is now an *a*-domain.

The most striking results obtained are shown in the center of Fig. 6. Since two *c*-domains have been introduced into the *a*-domain the reflection of the bright odd-shaped region (*g*) on the extreme right through the *a*-domains cannot take place in one step. The reflection is accomplished in 3 steps and the sum of the 3 dark images (*d*, *e*, *f*) in the center of Fig. 6 is the mirror image of the bright spot (*g*) and is equivalent to the dark area (*c*) from the first etching. This multiple reflection again exhibits the strong forward coupling of the electric dipoles in a beautiful way.

In addition to the reflection through *a*-domains as shown and discussed above, there also is a coupling through *c*-domains. In Fig. 5 we can see that the two sets of reflections are not independent of each other because the four odd-looking regions (*a*, *b*, *c*, *d*) look very similar to each other and are nicely aligned. It seems that the first and second pair of reflections are coupled through a *c*-domain. This is substantiated by the fact that there are other regions in Figs. 5 and 6 which show similar behavior. For example, at the upper edge of the photographs we can see two sets of cigar-shaped reflections (though the last reflection on the right is not too clear). How can we have reflection through a *c*-domain perpendicular to the dipole direction? The only way we can explain it is by the very strong piezoeffect in $BaTiO_3$. The crystal shown in Fig. 5 was originally a *c*-domain crystal. The *a*-domains which were introduced are wedge-shaped. Thus the crystal is probably strained in a complicated way but mostly sheared as shown in Fig. 7. The shear stress T_5 around an axis perpendicular to the paper (*y*-axis) induces a polarization P_i in the horizontal direction (*x*-axis) through the piezoelectric strain constant $d_{15}(P_x=d_{15}T_5)$. The sense of this induced polarization P_i depends on the sense of the spontaneous polarization P_s. For the 90° wall at the left we find that P_i points to the right for the projection on the domain wall of the odd-shaped antiparallel region (*b*) and points to the left for the normal surrounding region. If we go through the same argument for the 90° wall to the right we find that P_i would have the opposite directions. However, since P_i has to be continuous (electrostatic coupling) we must have P_i pointing to the right for the odd-shaped region (*c*) and to the left for the surrounding region. This can only be obtained if the spontaneous polarization P_s reverses. Thus the *c*-domain has to reverse its polarization in the middle, there must be a 180° wall. This is exactly what we observe, the 180° wall is very easily visible because one

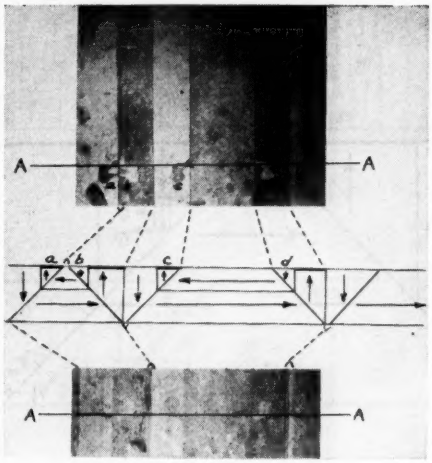


FIG. 5. Top and bottom view of a $BaTiO_3$ crystal showing domain reflections.

side of it etches much faster than the other according to the arguments given in Sec. III. Thus the two sets of reflections are connected through the strong piezoeffect in $BaTiO_3$. This coupling, however, requires a 180° wall in the *c*-domain. It can be seen that it is not as strong as the direct head to tail coupling because the two sets of reflections do not quite look the same, whereas the direct coupling gives perfect mirror images. Even so, it seems very remarkable that such strong coupling can be obtained through a *c*-domain with the help of a strong piezoeffect and electrostatic forces.

V. ETCH PATTERN OF TWINNED CRYSTALS

An interesting etch pattern of a twinned crystal is shown in Fig. 8. It shows a *c*-domain crystal containing

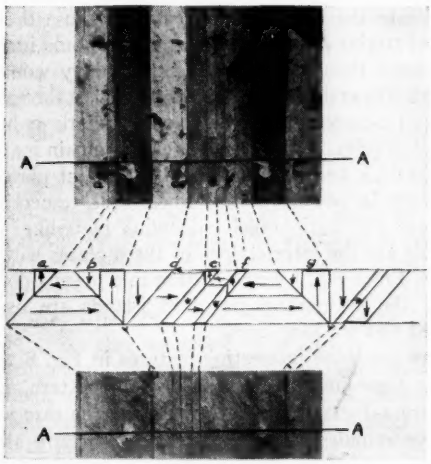


FIG. 6. Top and bottom view of a crystal showing a multiple domain reflection.

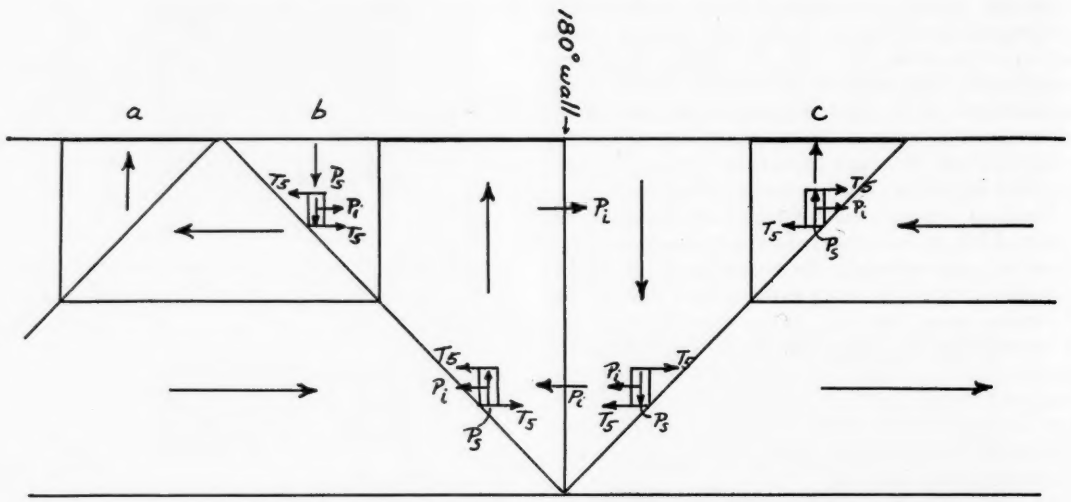


FIG. 7. Diagram illustrating the transverse coupling through a *c*-domain.

a few *a*-domains as can be seen very easily using transmitted polarized light [Fig. 8(b)]. The surface of the crystal, however, looks very different. It is shown in Fig. 8(a) using reflected light. The surface contains very many dark and bright lines which we call "channels." The picture looks very similar to Fig. 4 which was obtained optically by straining the crystal.[1] One is therefore tempted to believe that the etched crystal shown in Fig. 8 is a *c*-domain crystal with some *a*-domains, but covered with a layer of *a*-domain material at the surface. That is, it seems, that the polarization of the surface is parallel to the surface though the crystal is mostly a *c*-domain crystal. This surface polarization would be the electric analogue to the domains of closure in ferromagnetism.[2] However, we do not believe that the surface is polarized in a way different to the inside of the crystal because the mechanical strains would be very large and would increase the energy, though the electrostatic energy would be reduced. The etching experiment shows that the surface is not an *a*-domain because the etch is dark or bright and not greyish. Thus the surface is *c*-domain material and the dark and bright channels represent plate-like domains polarized upwards or downwards, respectively. Why are the antiparallel *c*-domains plate-like? The channels are the intersections of these plates with the surface. From Fig. 1 we know that in a poled *c*-domain crystal the antiparallel polarized islands are usually rounded and not long lines.

There are three interesting features in Fig. 8 which give us some insight into the domain pattern. First, there are *a*-domains present whereas in the case where we have rounded islands, as in Fig. 1, there are no *a*-domains. Actually there are more *a*-domains in this crystal than one observes in Fig. 8(b). All the sharp lines (actually they are double lines) on the surface [Fig. 8(a)] are the edges where *a*-domains intersect the surface. Only the thicker *a*-domains can be seen in transmitted light as the usual bright bands [Fig. 8(b)]. The thickness of the very thick *a*-domains can be measured directly by measuring the spacing of the double lines in Fig. 8(a) (center right, for example). The second feature is that the channels are always perpendicular to the aforementioned *a*-domain edges. If the channels coming from perpendicular sets of *a*-domains meet they try to join each other and form 45° lines. Third, the channels are more perfect (evenly spaced and parallel) the more *a*-domains we find in that particular region. Based on these 3 observations we propose that the channels, though they are *c*-domains, are caused by the *a*-domains. As a matter of fact, we have never observed channels on crystals where there were no *a*-domains and whenever we have a number of *a*-domains we can observe channels when the crystals are etched.

In order to understand the channels we assume a *c*-domain crystal with two widely-separated *a*-domains, as shown in Fig. 9(a). Since the two *a*-domains are far apart there will be no relation between their domain patterns. However, if the *a*-domains are closely spaced (overlap) as shown in Fig. 9(b) the following will happen. Let us assume that domain No. 1 consists of two parts, then only the lower part of domain No. 2 will be polarized to the left (head to tail coupling). But since the upper part of domain No. 1 is polarized to the left and since domain No. 2 is close to No. 1 we will find that also the upper part of No. 2 will be polarized to the left (heavy arrow) because of electrostatic forces. This is similar to the electrostatic coupling through a

[2] Williams, Bozorth, and Shockley, Phys. Rev. 75, 155 (1949).

c-domain as explained in Fig. 7. Now, the whole a-domain No. 2 is polarized to the left and thus the region above and below it has to be polarized upwards (heavy arrows). The same argument can be made for domain No. 3. Thus the whole top surface is polarized upwards except the intercepts with the a-domains, which are polarized to the left. Figure 9(c) shows the crystal viewed from the top and shows the plate-like domain. The sign of polarization of these plates alternates. The dark and bright bands are planes polarized upwards or downwards, respectively. The width of these channels is of the order of 10^{-4} cm or smaller (near wavelength of light) which agrees very well with the values obtained in Figs. 3 and 4 before.[1] Following the arguments given in reference 1 [optimal width of nuclei and weak sidewise coupling] we conclude that the width of 10^{-4} cm is determined by the nucleation

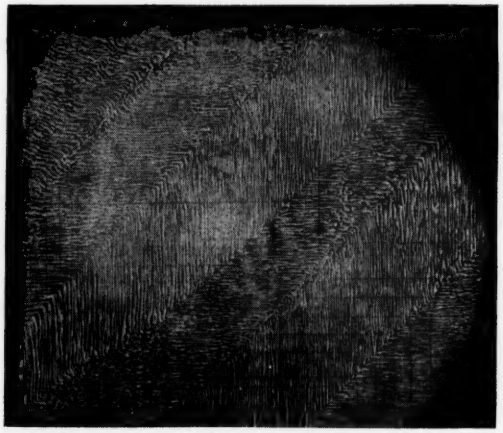


FIG. 8(a). Surface of a twinned $BaTiO_3$ crystal observed with reflected light.

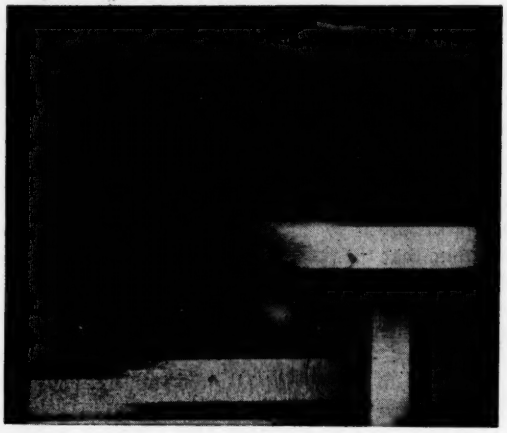


FIG. 8(b). Same crystal observed with transmitted polarized light.

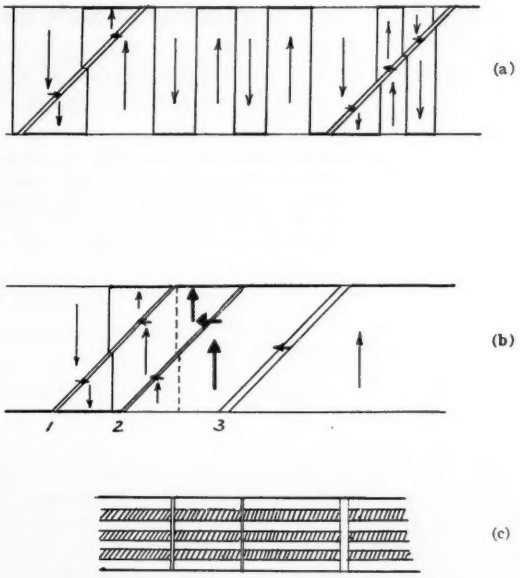


FIG. 9. Diagram illustrating the electrostatic coupling controlling the channel development.

and growth of the new domains. Of course, if we use a dc-poled crystal we never find channels. We only find channels, as shown in Fig. 8, when the crystal is ac-poled in such a way that about half of the dipoles point in one direction and the other half in the opposite direction. In this case, the nucleation can determine the size. In order to find channel etch patterns, we have to use ac-poled crystals where there are a-domains which are closely spaced so that the electrostatic coupling through the c-domain is large. If the a-domains are widely spaced, the interaction is small and the channels lose their sharpness and spread out in a disordered way. If wedge-shaped a-domains are present, as in Fig. 5, the channel etch pattern cannot be obtained.

VI. CONCLUSIONS AND ACKNOWLEDGMENTS

The etching technique, as described in the foregoing, is found to be a very useful tool for the investigation of the dipoles and domains in ferroelectric $BaTiO_3$ crystals. All the results obtained with optical methods can be verified and often clarified. The existence of the very strong forward coupling of the dipoles has been clearly demonstrated.

We would like to thank the following people who helped us on one or another stage of our work. Many problems were discussed with P. W. Anderson, S. S. Flaschen, and Mrs. E. A. Woods. E. M. Kelly, J. P. Remeika, and E. M. Walters helped us with the preparation of the samples.

Mobility of Slow Electrons in Polar Crystals*

FRANCIS E. LOW AND DAVID PINES†
Department of Physics, University of Illinois, Urbana, Illinois
(Received January 10, 1955)

An intermediate-coupling calculation of polaron-phonon scattering has been carried out by expressing the scattering amplitude directly in terms of matrix elements between exact eigenstates of the Hamiltonian corresponding to the initial and final states of the scattered particle. The basic approximation made is in the use of the rather accurate polaron wave functions of Lee, Low, and Pines for these initial and final states. The result for the mobility, which is only valid in the low-temperature region ($T \ll \theta$) is

$$\mu = \frac{1}{2\alpha\omega}\left(\frac{e}{m}\right)\frac{1}{[1+(\alpha/6)]^2} f(\alpha) \exp(\omega/\kappa T),$$

where α is the coupling constant for the electron-phonon interaction and $f(\alpha)$ is a slowly varying quantity of order 5/4 for $3<\alpha<6$.

I

A CONSIDERABLE amount of attention has been devoted in the past few years to the problem of a slow electron (or polaron) moving in the conduction band of a polar crystal.[1] The interaction between the polaron and the longitudinal optical mode of the crystal lattice vibrations is a strong one, for as the electron moves through the crystal its Coulomb field displaces the positive and negative ions with respect to one another; the resultant polarization charge density in turn appreciably distorts the motion of the electron. The strength of this mutual interaction can be characterized by a constant, α, which is of the rough order of the fourth root of the ratio of the atomic mass to the electron mass, and varies between 3 and 6 for most polar crystals. It is clear that for an interaction of this strength the usual perturbation theoretic methods of calculation (in which α is assumed to be much smaller than unity) are inadequate, and more powerful methods must be introduced. It is this "strong-coupling" aspect of the problem which has aroused so much interest on the part of the theoretical physicist.

The energy of a polaron of low momentum may be characterized by two parameters, E_0 and m^*. E_0 is the minimum energy required to raise an electron into the conduction band and m^* is the polaron effective mass. For the coupling constants of the size one finds in actual polar crystals, these quantities have been calculated quite accurately by the authors and Lee, by using a variational approach which is closely related to the intermediate-coupling method introduced by Tomanaga in meson theory.[2] In this method of calculation it was first assumed that all virtual phonons in the polaron cloud are emitted into the same state, characterized by a momentum distribution $f(k)$. This simple assumption makes possible a solution in closed form of the single polaron problem. The next approximation consists in allowing two kinds of virtual phonon states. Using this, Lee and Pines[3] have shown that corrections to the simple scheme described above were quite small for $\alpha<6$, from which we conclude that the LLP wave functions describe the polaron rather well in practice.

Unfortunately no direct experimental check on these calculations is at present available. The reason for this is that m^* and E_0 depend upon the "ionization" energy, I, and crystal mass, m, which characterize the behavior of a fictitious electron interacting with the static crystalline field (but not with the lattice vibrations). These are in turn unknown. The situation may be somewhat improved by an accurate calculation of the mobility, μ, of a polaron moving in a weak electric field. (μ is defined as the average polaron velocity per unit field strength.) A comparison of theory with experiment will then yield an empirical value for the crystal mass m. Finally, if it proves possible to make an experimental measurement of m^* in a cyclotron resonance experiment we should have the desired comparison between theory and experiment.

In this paper we carry out a calculation of the polaron mobility which we anticipate will be reasonably accurate in the "intermediate coupling" region of $\alpha<6$. Our procedure is to calculate $1/\tau$, the probability per unit time that the polaron be scattered by the thermal phonons. For sufficiently low temperatures, where the scattering is isotropic and elastic, the mobility is given by

$$\mu = e\tau/m^*.$$

It is, of course, somewhat difficult to calculate scattering in a strong-coupling situation. Our approach is based

* This work has been supported in part by the Office of Ordnance Research, U. S. Army.
† Now at Palmer Physical Laboratory, Princeton University, Princeton, New Jersey.
[1] H. Fröhlich, Advances in Phys. 3, 325 (1954), where a list of of references to earlier work may be found. We should also like to call attention to a recent paper on the theory of the polaron energy levels by R. P. Feynman, Phys. Rev. 97, 660 (1955).
[2] Lee, Low, and Pines, Phys. Rev. 90, 297 (1953), hereafter referred to as LLP. A method of approach substantially equivalent to that of LLP was developed independently by M. Gurari, Phil. Mag. 44, 329 (1953) and S. Zienau, British Electrical Research Association Report ERA—L/T 299, 1953 (unpublished), and by T. Yokota, Busseiron-Jenkyu 64, 148 (1953).
[3] T. D. Lee and D. Pines, Phys. Rev. 92, 883 (1953).

on the hope that we can overcome many of these difficulties by using the LLP wave functions to describe the initial and final states of the polaron. We might then expect that in so doing we will have taken into account a considerable fraction of the strong electron-phonon interaction, so that the effective coupling for the scattering of a polaron by a phonon is correspondingly reduced. However, the conventional method of treating scattering problems in field theory[4] expresses the scattering amplitude in terms of matrix elements between *bare* particle states. For this reason we make use of a different formalism recently derived by one of us[5] which expresses the scattering amplitude directly in terms of matrix elements between exact eigenstates of the Hamiltonian corresponding to the initial and final states of the scattered particle.

II

We adopt the Hamiltonian and notation of LLP, in which[6]

$$H=\sum_k a_k^* a_k\omega+\sum_k\{V_k a_k e^{i\mathbf{k}\cdot\mathbf{r}}+V_k^* a_k^* e^{-i\mathbf{k}\cdot\mathbf{r}}\}-\nabla^2/2m, \tag{1}$$

where

$$V_k=-\frac{\omega i}{k}\frac{1}{(2m\omega)^{\frac{1}{4}}}(4\pi\alpha)^{\frac{1}{2}}, \tag{2}$$

and

$$\alpha=\left(\frac{me^4}{\omega}\right)^{\frac{1}{2}}\left(\frac{1}{n^2}-\frac{1}{\epsilon_0}\right). \tag{3}$$

The dimensionless number α plays the role of a coupling constant for the lattice-electron interaction, a_k and a_k^* are phonon destruction and creation operators which satisfy the commutation relation $[a_k,a_{k'}^*]=\delta_{kk'}$, n and ϵ_0 are the optical index of refraction and the static dielectric constant of the crystal. In using the Hamiltonian we have thus assumed that the effect of the static periodic lattice field may be included in the crystal mass m.

According to LLP, the stationary state of the Hamiltonian (1) corresponding to a free polaron of momentum P_0 and energy $P_0^2/2m^*$ is well approximated by

$$\Psi(\mathbf{P}_0)=\exp\{i(\mathbf{P}_0-\textstyle\sum_k a_k^* a_k\mathbf{k})\cdot\mathbf{r}\}U(\mathbf{P}_0)\Phi_0, \tag{4}$$

where Φ_0 is the free vacuum state defined by

$$a_k\Phi_0=0, \quad (\Phi_0,\Phi_0)=1, \tag{5}$$

and where

$$U(\mathbf{P})=\exp\{\textstyle\sum_k[a_k^* f(\mathbf{k},\mathbf{P})-a_k f^*(\mathbf{k},\mathbf{P})]. \tag{6}$$

Here,

$$f(\mathbf{k},\mathbf{P})=-\frac{V_k^*}{\omega+(k^2/2m)-(\mathbf{k}\cdot\mathbf{P}/m)}. \tag{7}$$

It is shown in reference (5) that the S matrix element for the scattering of a polaron in momentum state $\mathbf{P}_0$ by a phonon of momentum $\mathbf{k}_0$ into a state specified by the momenta $\mathbf{P}_f$ and $\mathbf{k}_f$ is given by

$$S=1-2\pi i\delta(E_f-E_0)R, \tag{8}$$

where

$$R=V_{k_0}V_{k_f}^*\int d\mathbf{r}\Big\{\Big(\Psi_f(\mathbf{r}),e^{-i\mathbf{k}_0\cdot\mathbf{r}}\frac{1}{(H-E_f-\omega)-i\epsilon}\times e^{i\mathbf{k}_f\cdot\mathbf{r}}\Psi_0(\mathbf{r})\Big)+\Big(\Psi_f(\mathbf{r}),e^{i\mathbf{k}_0\cdot\mathbf{r}}\times\frac{1}{(H+\omega-E_0)-i\epsilon}e^{-i\mathbf{k}_f\cdot\mathbf{r}}\Psi_0(\mathbf{r})\Big)\Big\}. \tag{9}$$

Here $\Psi_0(\mathbf{r})$ and $\Psi_f(\mathbf{r})$, are the exact initial and final eigenstates of the polaron, and ϵ is an infinitesimal positive constant introduced to ensure that our wave functions properly correspond to outgoing waves in a coordinate space representation. E_0 and E_f are the initial and final polaron energies; e.g.,[7]

$$E_0=P_0^2/2m^*, \quad E_f=P_f^2/2m^*,$$

for sufficiently small polaron energies and H is the total Hamiltonian defined by Eq. (1). It should be emphasized that Eq. (9) is exact. We now, however, will make our first approximation, which consists in using the solution (4) of LLP for Ψ_0 and Ψ_f. As pointed out by LLP, these wave functions are quite accurate for low momenta such that

$$(P^2/2m^*)\ll\omega,$$

but should not be taken seriously when $P^2/2m^*$ becomes comparable with ω. Thus our work will only be valid for temperatures such that $\kappa T\ll\omega$ or $T<\theta_D$, the Debye temperature for the longitudinal optical mode.

In the following we shall therefore make a virtue of necessity by actually considering only the $P\to 0$ limit, a procedure which has the advantage of enormously simplifying our calculations. This limit, however, must be taken with considerable care. Since in a normal scattering process the cross section, σ, approaches a finite limit as $P\to 0$, the transition rate, $w=P\sigma$, will approach zero with P unless something peculiar happens to the scattering matrix element. To be more precise,

$$w=2\pi\rho_E|M|^2d\Omega, \tag{10}$$

where ρ_E is the density in energy of final states and M is an appropriate matrix element which is simply related to the R matrix of Eq. (8); the point is that $\rho_E\to 0$ as $P\to 0$, and we will therefore only find a finite transition rate when M has a singularity at $P=0$. Such a singularity will, in fact, occur when the polaron is scattered by a phonon whose momentum is such that it can be

[4] F. Dyson, Phys. Rev. **75**, 1736 (1949).
[5] F. Low, Phys. Rev. **97**, 1392 (1955).
[6] We work in units in which $\hbar$ and the volume of the crystal are unity.
[7] We are measuring polaron energies with respect to the constant binding energy $E_0=-(\alpha\omega+I)$.

absorbed by the polaron with approximate conservation of energy. Since the width of this "excited" polaron state (polaron+absorbed phonon) approaches zero with P, the subsequent re-emission of the phonon will lead to a resonant scattering matrix-element of the form

$$M(k_0)\sim\frac{B}{(k_0^2-k_r^2)+iAP},$$

where k_r is the resonant phonon wave-vector. It follows that

$$w\sim\frac{CP}{(k_0^2-k_r^2)^2+A^2P^2}\sim D\delta(k_0^2-k_r^2)$$

in the limit of $\mathbf{P}\to 0$. The integration of w over all incident phonon wave-vectors k_0 then yields a finite result in this limit. Thus in what follows we shall let $\mathbf{P}\to 0$ almost everywhere, but we shall be careful to keep those $\mathbf{P}$ dependent terms which lead to the $\mathbf{P}$-dependent width in the denominator of $M(k_0)$.

If we now substitute (4) and (6) into our expression (9) for the R matrix, we may explicitly carry out the integration over $\mathbf{r}$ by commuting all $\mathbf{r}$ dependent factors through the energy denominators from right to left. In the limit of $\mathbf{P}\to 0$ the foregoing discussion makes it clear that the second term of (9), with denominator $(H+\omega-E_0)-i\epsilon$, will not contribute to our transition rate, for E_0 is the lowest eigenvalue of H, and therefore $H+\omega-E_0\gtrless\omega$ and can never vanish. Thus we obtain

$$R=(2\pi)^3\delta(\mathbf{P}_f+\mathbf{k}_f-\mathbf{P}_0-\mathbf{k}_0)V_{k_0}V_{k_f}^*\left[U(P_f)\Phi_0,\frac{1}{\bar{H}[\mathbf{P}_0-\sum_k(a_k^*a_k\mathbf{k})+\mathbf{k}_0]-E-\omega-i\epsilon}U(P_0)\Phi_0\right], \quad (11)$$

where $E=E_0=E_f$ and

$$\bar{H}(\mathbf{P})=(P^2/2m)+\sum_k V_k a_k+V_k^* a_k^*+\sum_k a_k^* a_k\omega. \quad (12)$$

We further commute $U(\mathbf{P}_0)$ through to the left in (10), and obtain for the matrix element:

$$R=(2\pi)^3\delta(\mathbf{P}_f+\mathbf{k}_f-\mathbf{P}_0-\mathbf{k}_0)V_{k_0}V_{k_f}^*\times\left(\Phi_0,\frac{1}{H_0+H_1-(P_0^2/2m^*)-\omega-i\epsilon}\Phi_0\right). \quad (13)$$

We have set $U^{-1}\bar{H}U-E=H_0+H_1-P_0^2/2m^*$ and $U^{-1}(\mathbf{P}_f)U(\mathbf{P}_0)=1$, this latter identity applying in the zero momentum limit. H_0 and H_1 are given by

$$H_0=\frac{P_0^2}{2m^*}+\left[\sum_k a_k^*a_k\left(\omega-\frac{\mathbf{k}\cdot\mathbf{k}_0}{m}-\frac{\mathbf{k}\cdot\mathbf{P}_0}{m^*}\right)\right]+\frac{k_0^2}{2m}+\frac{\mathbf{k}_0\cdot\mathbf{P}_0}{m^*}+\frac{(\sum_k a_k^*a_k\mathbf{k})^2}{2m}, \quad (14a)$$

$$H_1=\sum_{kk'}\frac{\mathbf{k}\cdot\mathbf{k}'}{2m}\{f^*(\mathbf{k})f(\mathbf{k}')a_k a_{k'}+2a_k^*a_{k'}f(\mathbf{k})f^*(\mathbf{k}')+a_{k'}^*a_k^*f(\mathbf{k})f(\mathbf{k}')+2a_k^*a_k a_{k'}f^*(\mathbf{k}')+2a_{k'}^*a_k^*a_k f(\mathbf{k})\}-\sum_k\frac{\mathbf{k}\cdot\mathbf{k}_0}{m}\{f(\mathbf{k})a_k^*+f^*(\mathbf{k})a_k\}. \quad (14b)$$

Here $f(\mathbf{k})$ is given by (7) and we have suppressed the index P_0. We may then write

$$S=1-(2\pi)^4i\delta(E_f-E_0)\delta(\mathbf{P}_f+\mathbf{k}_f-\mathbf{P}_0-\mathbf{k}_0)M, \quad (15)$$

where

$$M=|V_{k_0}|^2\left(\Phi_0,\frac{1}{H_0+H_1-(P^2/2m^*)-\omega-i\epsilon}\Phi_0\right), \quad (16)$$

and we have used the fact that $k_0=k_f$ in the zero polaron momentum limit. The transition rate is given by (10), so that our remaining task is the evaluation of the matrix element (16).

We do this by introducing

$$c\Psi=\frac{1}{H_0+H_1-(P^2/2m^*)-\omega-i\epsilon}\Phi_0, \quad (17)$$

where

$$c=\frac{1}{(k_0^2/2m)+(\mathbf{k}_0\cdot\mathbf{P}_0/m^*)-\omega}.$$

Ψ then satisfies the familiar equation

$$\Psi=\Phi_0-\frac{1}{H_0-(P^2/2m^*)-\omega-i\epsilon}H_1\Psi. \quad (18)$$

Ψ may be expanded in states of zero, one, two phonons, etc.; i.e.,

$$\Psi=u\Phi_0+\sum_k v_k a_k^*\Phi_0=\sum_{kk'}w_{kk'}a_k^*a_{k'}^*\Phi_0+\cdots. \quad (19)$$

We then find that our desired matrix element is

$$M=cu|V_{k_0}|^2. \quad (20)$$

We will keep only the zero- and one-phonon terms in our expansion (19). This approximation is essentially equivalent to that of LLP in that it involves neglect of just those terms in H_1 which LLP neglected in calculating the ground state wave function and energy. Indeed, it would be inconsistent with our use of the LLP wave functions were we to attempt a more accurate solution of (18), and it would in fact be quite possible for such a procedure to lead to less accurate results than those we obtain below.

When we substitute the expansion (19) into (18), keeping only zero- and one-phonon terms, we obtain

the following set of coupled equations:

$$u=1+c\sum_k v_k(\mathbf{k}\cdot\mathbf{k}_0/m)f^*(\mathbf{k}), \tag{21a}$$

$$v_k=\frac{u(\mathbf{k}\cdot\mathbf{k}_0/m)f(\mathbf{k})}{\Delta E}-\frac{f(\mathbf{k})\mathbf{k}}{\Delta E}\cdot\sum_{k'}\frac{\mathbf{k}'}{m}v_{k'}f^*(\mathbf{k}'), \tag{21b}$$

where

$$\Delta E=[(\mathbf{k}_0-\mathbf{k})^2/2m]+[(\mathbf{k}_0-\mathbf{k})\cdot\mathbf{P}_0/m^*]-i\epsilon. \tag{22}$$

Since the dependence of v_k on its variable k is explicitly given in (21b), Eqs. (21a) and (21b) are essentially algebraic in nature, and may be solved exactly. We find,

$$cu=\left[\frac{k_0^2}{2m}\left(\frac{1-I}{1+I}\right)-\omega\right]^{-1}, \tag{23}$$

where

$$I=\frac{1}{m}\sum_k\frac{(\mathbf{k}\cdot\mathbf{k}_0)^2|f(\mathbf{k})|^2}{k_0^2\Delta E(\mathbf{k},\mathbf{P}_0)}. \tag{24}$$

In obtaining (23) and (24) we have set $P_0=0$ everywhere except in the expression (22) for ΔE. This is in keeping with our announced program of letting P go to zero everywhere but in the width. We may replace the summation in (24) by an integral over k. We recall that $1/(a-i\epsilon)$ is a conventional notation for $\mathcal{P}(1/a)+i\pi\delta(a)$, where $\mathcal{P}$ denotes the Cauchy principal value, and find

$$I=I_1+iI_2, \tag{25}$$

with

$$I_1=2\int\frac{d\mathbf{k}}{(2\pi)^3}\frac{(\mathbf{k}\cdot\mathbf{k}_0)^2|f(\mathbf{k})|^2}{k_0^2(\mathbf{k}-\mathbf{k}_0)^2}+O(P), \tag{26}$$

and

$$I_2=\frac{2P_0\alpha\omega^2}{(2m\omega)^{\frac{1}{2}}[\omega+(k_0^2/2m)]^2}\left(\frac{m}{m^*}\right). \tag{27}$$

The integral I_1 is elementary, and is most easily carried out by integrating first over k, and second over angles. It is

$$I_1=\alpha F[k_0/(2m\omega)^{\frac{1}{2}}], \tag{28}$$

where

$$F(x)=\pi x\left[\frac{1}{1+x^2+|1-x^2|}\right]^2-\frac{1+x^2}{2x^2}\left\{\frac{1+x^4}{(1+x^2)^2}-\frac{(x^2-1)}{2x}\tan^{-1}\frac{2x}{x^2-1}\right\}. \tag{29}$$

We now substitute Eq. (20) and (23) into our expression (10) for the lifetime. On integrating this over all incident phonons k_0 and the solid angle $d\Omega_{Pf}$, we obtain the mean lifetime for scattering of a polaron,

$$\frac{1}{\tau}=\frac{m^*P_0}{8\pi^4}\int\frac{d\mathbf{k}_0|V_{k_0}|^4e^{-\omega/\kappa T}}{|\omega-(k_0^2/2m)([1-I(k_0)]/[1+I(k_0)])|^2} \tag{30}$$

where we have used the fact that the mean number of phonons of wave-vector k_0 present at temperature $T\cong e^{-\omega/\kappa T}$. For low polaron momentum we may re-write

$$\frac{1-I}{1+I}=\frac{1-I_1}{1+I_1}-\frac{2iI_2}{(1+I_1)^2},$$

so that

$$\frac{1}{\tau}=\frac{m^*P_0e^{-\omega/\kappa T}}{8\pi^4}\int\frac{4\pi k_0^2dk_0|V(k_0)|^4}{\left\{\left[\omega-\frac{k_0^2}{2m}\left(\frac{1-I_1}{1+I_1}\right)\right]^2+\frac{k_0^4}{m^2}\frac{I_2^2}{(1+I_0)^4}\right\}}. \tag{31}$$

We set $I_2=AP_0$, and apply the identity

$$\operatorname*{Lim}_{P\to0}\frac{P}{B^2+C^2P^2}=\frac{\pi}{C}\delta(B) \tag{32}$$

to obtain

$$\frac{1}{\tau}=\frac{m^*e^{-\omega/\kappa T}}{2\pi^2}\int\times\frac{dk_0k_0^2|V(k_0)|^4\delta[\omega-(k_0^2/2m)(1-I_1)/(1+I_1)]}{[Ak_0^2/m(1+I_1)^2]}. \tag{33}$$

Finally we recall that

$$\int dk_0f(k_0)\delta[\omega-g(k_0)]=f(k_r)\Big/\left(\frac{\partial g}{\partial k_0}\right)_{k_0=k_r},$$

where k_r is the root of $g(k_r)=\omega$, and find, after some algebraic manipulation,

$$\frac{1}{\tau}=\left(\frac{m^*}{m}\right)^2\frac{2\alpha\omega e^{-\omega/\kappa T}}{f(\alpha)}, \tag{34}$$

where

$$f(\alpha)=\left[\frac{1}{x_r}-\frac{(1+x_r^2)^2}{4x_r^2}\left(\frac{\partial I_1}{\partial x}\right)_{x=x_r}\right]. \tag{35}$$

I_1 is given by (28) and x_r is the root of

$$x_r^2=\frac{1+I_1(x_r)}{1-I_1(x_r)}. \tag{36}$$

The mobility μ is then[8]

$$\mu=\frac{1}{2\alpha\omega}\frac{e}{m}\left(\frac{m}{m^*}\right)^3f(\alpha)e^{\omega/\kappa T}. \tag{37}$$

[8] This differs from our previously announced result [Phys. Rev. **91**, 193 (1953)] by a factor of (m/m^*). This error we attribute to a too cavalier treatment of the limit $P\to0$. It might be added that our previous result was obtained by using a somewhat different scattering formalism which turns out not to be as well suited to our problem as the one which we apply here.

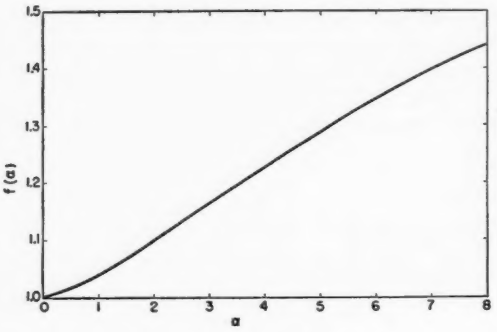


FIG. 1. $f(\alpha)$ *vs* α.

The function $f(\alpha)$ is a rather slowly varying function of α, as may be seen in Fig. 1. In the perturbation theory limit, $I_1=0$, $x_r=1$, and $f(\alpha)=1$, so that we get the correct weak-coupling value for the mobility.[9]

[9] See reference 1 or footnote 4 of reference 8 for the correction of the earlier published values.

It is worth noting that although the function $f(\alpha)$ is very nearly unity for coupling strengths encountered in polar crystals, the dependence of μ on $(m/m^*)^3 = (1+\alpha/6)^{-3}$ leads to the conclusion that the crystal mass m is rather smaller than one would estimate from mobility experiments using the perturbation-theoretic expression for free mobility. Thus for the experiments of Breckenridge and Hosler[10] on TiO_2, we find $m \cong 9m_e$ in place of the considerably higher value $m \cong 29m_e$ obtained by using the correct perturbation theoretic value for their sample 750.

We will be surprised if our expression for the mobility turns out to be wrong by more than 30 percent in the temperature and coupling region we have considered.

We should like to thank Professor John Bardeen and Professor Tsung-Dao Lee for stimulating discussions on this subject and related problems.

[10] R. G. Breckenridge and W. R. Hosler, Phys. Rev. **91**, 793 (1953).

PHYSICAL REVIEW VOLUME 98, NUMBER 2 APRIL 15, 1955

Neutron Irradiation Effects in Cu and Al at 80°K*

A. W. McREYNOLDS, W. AUGUSTYNIAK, MARILYN McKEOWN, *Brookhaven National Laboratory, Upton, New York*

AND

D. B. ROSENBLATT, *Pitman-Dunn Laboratories, Frankford Arsenal, Philadelphia, Pennsylvania*

(Received December 31, 1954)

A cryostat is described for pile neutron irradiation in the BNL reactor of samples at liquid nitrogen temperature. One-month irradiation resulted in resistivity increases of 20 and 33 percent for Cu and Al, respectively, at 80°K, and increases by about a factor of 5 in critical shear stress for both. Subsequent studies of the kinetics of recovery processes were made by a number of isothermal annealing curves at various temperatures from −80°C to 500°C. In copper, two-thirds of the resistivity effect anneals out from −80° to 20°C, activation energy ~0.6 ev, with no decrease in critical shear stress effect. Both effects then anneal together near 320°C, with energy 2 ev. In aluminum, both effects completely anneal out together near −60°, with 0.55 ev energy. Data are correlated with diffusion activation energy for lattice defects.

I. INTRODUCTION

A FAST nuclear particle passing through a crystal lattice ejects a number of atoms from their sites by collision. The effects of the resulting lattice defects on the physical properties have been extensively studied in the past few years. The defects may reasonably be assumed to be initially primarily vacancies, interstitials or clusters of these. In general however the nature of the defects, their effect on physical properties, or the kinetics of their motion and recombination are not yet well established. The present studies were directed toward further information on these three points, through observations on the kinetics of the annealing process by which radiation effects are healed.

It has been shown in prior work that although semiconducting and insulating crystals show large effects, particularly in electrical and optical properties, from pile irradiation, the effects in metals are much smaller.[1] This is partly accounted for by the lesser sensitivity of metals to small numbers of defects. It was also assumed, as is now verified by these and other experiments,[2–4] that a large fraction or all of the radiation induced lattice defects may anneal out rapidly even at room temperature, although higher temperatures are required for diffusion and for annealing of dislocations. In order to "freeze in" the radiation effects apparatus

* Under the auspices of the U. S. Atomic Energy Commission

[1] Review articles of earlier work are: D. S. Billington and S. Siegel, Metal Progr. **58**, 847–52 (1950); J. C. Slater, J. Appl. Phys. **22**, 237–56 (1951); F. Seitz, Discussions Faraday Soc. **5**, 271–82 (1949); G. J. Dienes, Ann. Rev. Nuc. Sci. **2**, 187 (1953).

[2] Albert W. Overhauser, Phys. Rev. **90**, 393 (1953).

[3] Marx, Cooper, and Henderson, Phys. Rev. **88**, 106 (1952).

[4] R. R. Eggleston, Acta Metallurgica **I**, 679 (1953).

has been developed for keeping samples at the center of the BNL nuclear reactor continuously at liquid nitrogen temperature for long periods of time. At this temperature even such defects as interstitial atoms and lattice vacancies should be relatively stable. On the other hand, although the macroscopic temperature is low, a large amount of ionization energy is dissipated in the collision processes causing intense local heating in just the small region around the defects. It might be expected then that some immediate rearrangement would partly heal the defects. It has been indicated by Brooks[5] that most likely this local heat is dissipated through the electronic system too rapidly to be communicated to the nuclei as vibrational energy. This conclusion is substantiated by the present results of low temperature neutron irradiations and results of low temperature cyclotron deuteron bombardments,[2,3] both of which indicate that a significant amount of damage can be frozen in.

In order to correlate measurements on various physical properties, a large number of identically prepared wires of two face-centered cubic metals were simultaneously irradiated. Dimensions were chosen to be suitable for subsequent measurements of electrical resistivity, elastic plastic, and anelastic mechanical properties, as well as diffraction examinations of the structure.

As described below, of these properties only electrical resistivity and critical shear stress were found to be suitable indices of the degree of radiation damage to the lattice. Electrical resistivity is convenient because it should be expected to vary roughly linearly with the number of any particular kind of defect, exhibits moderately large effects, and can be easily and precisely measured under a variety of experimental conditions. Critical shear stress, although more difficult and less precise to measure, has the advantages that it is sensitive to smaller amounts of irradiation and exhibits even larger changes of the order of a few hundred percent. Theoretically it is somewhat difficult to interpret quantitatively, but offers the possibility of studying the interaction of interstitial atoms and vacancies with dislocations.

II. EXPERIMENTAL METHODS

Production and observation of fast-neutron radiation effects on the samples required their irradiation near the center of the reactor for a period of the order of one month while maintaining continuously a temperature near that of liquid nitrogen, and subsequent physical measurements while still at low temperature. Since development of a cryostat for the purpose involved mostly problems peculiar to reactors rather than the usual low-temperature techniques, general design considerations will be discussed briefly. The available radiation space in the BNL reactor is in the form of a 4-inch square hole extending through the shield and the graphite structure of the reactor. Coolant must be transferred from an external source horizontally along this hole to a suitably insulated container 18 feet from the outside wall. Further, a mechanism is required for removal of the samples in a manner not allowing appreciable temperature rise.

Restrictive conditions inherent in reactor irradiations are: (1) Access holes through the shield for coolant and samples must be shielded to prevent radiation leakage. (2) Activation during radiation prevents removal of the cryostat for alteration or repair and necessitates heavily shielded handling facilities for samples removed. (3) Construction materials must be limited to a few elements which are not strongly activated by irradiation and which have low enough absorption cross section not to deplete the neutron flux through the sample. (4) Further, the physical properties of the construction materials must not be seriously affected by radiation damage (i.e., all plastics, waxes or oils, rubber, etc., are excluded, but metals or other crystals and glass may be used). (5) Absorption of γ-ray and neutron flux generates heat directly in the sample and internal cryostat parts, so that the lower limit of heat input is determined rather by the mass of the system than by its insulation against thermal conduction from outside. (6) Since irradiations of the order of a month are required and it is essential that low temperature be uninterrupted during this period, an automatic control system for supply of coolant is necessary and must have small probability of operational failure.

Several alternative systems were considered, in particular, blowing in a continuous stream of cold N_2 gas from a liquid N_2 tank outside the reactor, or circulating helium gas between the cryostat and an external liquid N_2 tank by means of a compressor. Such methods of continuous circulation, however, require maintenance of low temperatures not only around the sample, but also along the entire 18-foot length of leads from the external apparatus. The consumption of liquid nitrogen is thus large because of the large area for thermal conduction and large mass of radiation absorbing material. It was therefore more satisfactory to have only a small volume around the samples kept cold by a small reservoir of liquid nitrogen inside the reactor, the reservoir being replenished by periodic additions of a few cubic centimeters of liquid. It was found that such small quantities of nitrogen could be projected by air pressure along a horizontal warm metal tube with loss of only about 30 percent between external tank and internal cryostat.

The "irradiation cryostat" used is shown in Fig. 1. The inner quartz Dewar flask was designed to have as small mass as practicable and, since it had to be horizontal, a small mouth to minimize air circulation by convection. It is enclosed in a thin aluminum tank to which are connected five aluminum tubes leading out through the reactor shield. One of these, of 1-inch

[5] H. Brooks (private communication).

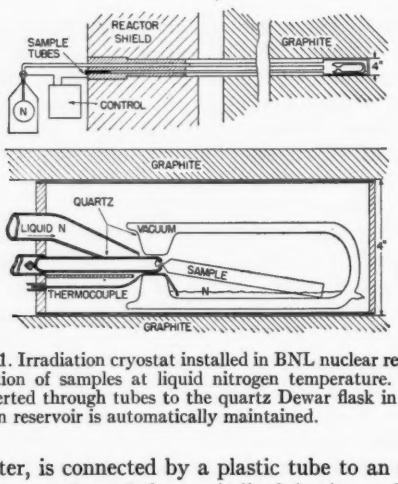


FIG. 1. Irradiation cryostat installed in BNL nuclear reactor for irradiation of samples at liquid nitrogen temperature. Samples are inserted through tubes to the quartz Dewar flask in which a nitrogen reservoir is automatically maintained.

diameter, is connected by a plastic tube to an outside container and used for periodic injections of liquid nitrogen. The other four parallel tubes, 0.400-inch inside diameter, permit insertion of four independent samples. The outer ends of these sample tubes are normally closed by shielding plugs allowing only an exit aperture for the nitrogen gas, so that the system is sealed against accumulation of moisture during prolonged operation. The sample containers are 6-inch long aluminum tubes, coupled at one end to a length of quartz tube. Because radiation heating of the sample necessitates good thermal contact with the coolant, the coupling is designed to allow one end of the aluminum tube to rest in the pool of nitrogen on the bottom. The quartz tube extends out of the cold region into the entrance tube. The sample containers can be inserted individually while the reactor is in operation by simply being pushed along the tube. For removal a small clip device is pushed in on the end of a steel wire to grip a knob on the end of the quartz tube to pull it out. The removal operation could be accomplished with a time of 10–15 seconds between departure of sample from the inner cryostat and its storage in a shielded flask of liquid nitrogen outside the reactor. This is rapid enough both to prevent appreciable warming of the sample during removal and to make radiation exposure of operating personnel negligibly small. After removal from the reactor, samples can be stored indefinitely under liquid nitrogen to allow decay of activation before making physical measurements. Automatic control of the cryostat is accomplished through a chromel-alumel thermocouple with junction located at the bottom of the quartz vacuum vessel and a Leeds and Northrup recording and controlling potentiometer. When temperature rises above a preset value, a control unit outside the pile applies compressed air pressure for a few seconds to the tank outside, injecting additional nitrogen. Since the neutron absorption cross section of nitrogen is small, but not negligible, it is desirable for maximum flux to keep only a small supply near the sample. Procedure for the two irradiations performed was therefore to inject nitrogen for a period of about six seconds, a quantity of about 100 cc. After this had boiled away and the inside temperature had risen to −150°C, the controller initiated the next injection cycle. Under normal operating conditions, the time for one complete cycle ranged from 4 to 10 minutes and consumption of liquid nitrogen was one to two liters per hour, depending on reactor power level, mass and nature of samples, and operation features of the filling mechanism. Rate of heat absorption of the cryostat was measured to be about three times as great with the reactor at full power as from thermal conduction alone when the reactor was shut down. A dummy sample containing a thermocouple was used to verify, however, that radiation heating did not cause the sample temperature to exceed that recorded from the controlling thermocouple. It should be mentioned that the cryostat apparatus described and shown in Fig. 1 has been succeeded by a second model. The present BNL facility, which has been in operation about two years, has the same general operating characteristics, but uses an all aluminum system with continuously pumped evacuated space between inner and outer vessels. This revised apparatus was developed subsequent to these experiments by R. W. Powell, J. Fleeman, and others.

The samples used in the measurements reported here were 99.99 percent pure copper and standard 2-S aluminum, both drawn into wire form. These materials were chosen primarily because there is a considerable body of information on their properties and secondarily because their nuclear properties, 13-hr half-life for Cu, 2.3-min half-life for Al, make possible physical measurements without shielding precautions after a relatively short storage period. The copper was drawn into wires of 0.028-in. and 0.013-in. diameter, vacuum-annealed for two hours at 450°C, and furnace cooled before irradiation. The aluminum was drawn into 0.032-in. wire and also annealed for two hours at 450°C. The aluminum sample containers were filled with a number of identical 5-inch lengths of the annealed wires for simultaneous irradiation. The first irradiation was for a period of 30 days with an integrated neutron flux of approximately 1.1×10^{19}. The second group of samples was irradiated for the same time, but is not directly comparable to the first since a temporary apparatus breakdown after 15 days allowed temperature to rise for a few minutes to a maximum of about −50°C, resulting in an undetermined amount of annealing.

Measurements of physical properties after irradiation were made without radiation shielding, and by techniques standard except for design of apparatus to allow installation of samples transferred at 80°K from the cryostat and to allow electrical and mechanical measurements on samples immersed in liquid nitrogen.

Electrical resistivity was measured by passing a current of a few amperes through the sample and comparing by a Leeds and Northrup type K potentiometer the potential drops across the sample, an unirradiated

control sample, and a standard resistor. For measurements with the 0.013-inch copper wire, current and potential leads were soldered to each end of five-inch wires, all but a few millimeters of the end being kept immersed in liquid nitrogen to prevent annealing by the soldering process. The 0.028-inch copper and aluminum samples were mounted in a jig with four knife-edge contacts as current and potential leads. All measurements were made at 80°K, with the sample immersed in liquid nitrogen, the control sample being immersed in the same bath to correct for slight temperature variations. Relative resistance measurements were accurate to 0.1 percent or better, but it was not attempted to measure dimensions of samples to this accuracy for determination of absolute resistivity. Apparatus is shown schematically in Fig. 2. Annealing was accomplished stepwise by rapid transfer of the sample holder to a second liquid bath held at the desired temperature and returned to the nitrogen bath for measurement. Since the samples reached bath temperatures in a time of the order of a second, annealing intervals could thus be as little as ten seconds. For temperatures −80°C to 50°C an alcohol bath was used, and for higher temperatures, oils, molten wax, or salt baths.

Mechanical properties were studied by means of a torsion pendulum as illustrated in Fig. 2. The pendulum was suspended from an upper fiber of 0.004-inch diameter wire, which had negligible rigidity compared to the samples. The sample wires were held in pin vises or spring clamps at each end, the arrangement being such that installation could be made under liquid nitrogen. Torque could be applied to the system by a galvanometer-type arrangement of a magnetic coil on the pendulum between the poles of a permanent magnet, and deflection was measured by a conventional optical system. Natural frequency of torsional oscillation provided a measure of elastic moduli, and decay of amplitude a measure of internal friction. Torsional stress-strain curves were obtained from current *vs* deflection data. Although plastic strain is nonuniform, flow beginning at the outer surface of the sample, the critical shear stress may be readily determined and is less dependent on sample geometry than in usual tensile tests. As in the case of the electrical resistivity, most measurements were made with the samples immersed in a flask of liquid nitrogen, and annealing between measurements was accomplished by rapid removal of the nitrogen flask and substitution of an alcohol bath at desired temperature. For annealing temperatures above 100°C, this method could not be used, and it was necessary to remove the sample for annealing in a wax or salt bath and replace it later. Since stress-strain curves were a destructive test, one sample was required for each run, and annealing was carried out before installation.

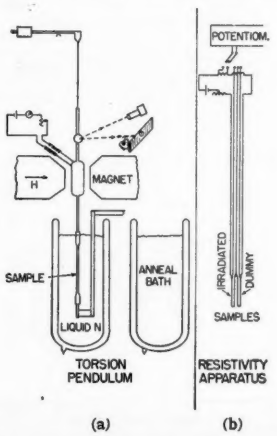


FIG. 2. (a) Torsion pendulum for low-temperature measurements of mechanical properties of irradiated wires. (b) Wire sample mounted for electrical resistivity measurements under nitrogen and annealing in liquid bath.

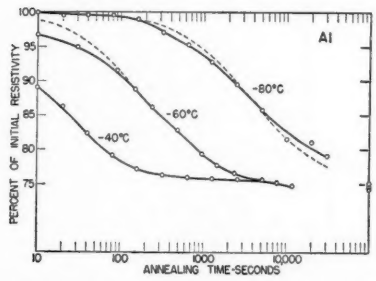


FIG. 3. Annealing of resistivity induced in Al wires by neutron irradiation. Annealing was in successive steps with all resistivity measurements made at 80°K. Initial resistivity refers to the irradiated state before any annealing. Dashed curve is of the form of Eq. (3).

III. RESULTS

Resistivity

Resistivity of the Al wires was measured in the irradiated condition and as a function of annealing time at temperatures of −40°, −50°, −60°, and −80°C. One sample was used for each temperature, and the annealing carried out in successive steps, with all measurements of resistivity being made at 80°K. Results are shown plotted in Fig. 3 as percent of the initial (irradiated) value. The corresponding value for fully annealed material could be determined more accurately by measurement after prolonged annealing than from data prior to irradiation, and is indicated in each case by the final points on the curve. If the assumption is made that the radiation induced excess resistivity depends linearly on the number of defects, and that these anneal out by a single process, the relation

$$dR/dt = -kR^n \tag{1}$$

should govern the isothermal annealing. This leads to

$$R = R_0 e^{-at} \quad \text{for} \quad n=1, \tag{2}$$

$$R = (at+b)^{-1} \quad \text{for} \quad n=2, \tag{3}$$

$$R = (at+b)^{-\frac{1}{2}} \quad \text{for} \quad n=3. \tag{4}$$

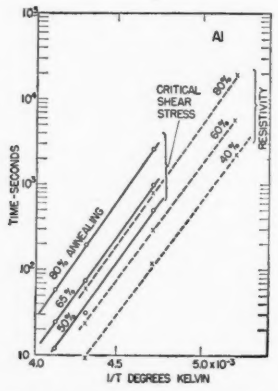


FIG. 4. Annealing data on Al wires from Figs. 3 and 9 plotted to determine heat of activation. Points represent the time to anneal out the indicated percent of the radiation induced excess resistivity and critical shear stress, respectively.

The data in Fig. 3 are found to be fit fairly well by the second of these relations, $n=2$, which is shown in the figure as the dashed curves, adjusted for fit at 50 percent annealing. Such dependence on the square of the defect concentration seems reasonable if the annealing process occurs by recombination of interstitial atoms and vacancies. Further, if the process is governed by a heat of activation, as would be expected if limited by diffusion rates, the rate constant a in Eqs. (1)–(3) should vary with temperature according to

$$a=a_0e^{-E/kt},$$

where E is the activation energy.

As pointed out by Parkins, Dienes, and Brown,[6] the heat of activation can be determined from a plot of $\ln t$ *versus* $1/T$, for equal degrees of annealing at different temperatures, regardless of the functional dependence of R on t and T. Such a plot is shown in Fig. 4, for stages where excess resistivity has decreased by 40, 60, and 80 percent of its original value. The linearity and parallelism of these three curves demonstrates that the annealing is governed by a single activation energy over the range studied. The value derived is

$$E=0.55 \text{ ev}.$$

Critical shear stress data included in the same figure will be discussed later.

The annealing process in Cu as shown in Figs. 5 and 6 is somewhat more complex and occurs in two or more distinct stages. For each temperature in the range -80°C to -20°C, the decrease of resistivity follows rather accurately from 10 to 10 000 seconds an equation of the form

$$\frac{dR}{d\ln t}=-\frac{1}{t}\frac{dR}{dt}=F(T), \qquad (5)$$

where $F(T)$ is a function of temperature but not time.

[6] Parkins, Dienes, and Brown, J. Appl. Phys. **22**, 1012 (1951).

Thus, the progress of annealing is much too rapid at short times to be consistent with any of Eqs. (1–4) as comparison with Fig. 4 indicates. This result is also in contrast with Overhauser's measurements on deuteron-irradiated copper, which were found to fit Eq. (2) with $n=2.5$. With increasing temperatures, the data show some indication of a second stage in the annealing process less dependent on temperature. Although annealing appears nearly completed at 1000 seconds at 20°C, it continues at about the same rate after 1000 seconds at 100°C. Relative heights of the curves for various temperatures in Fig. 5 cannot be compared significantly, however, since slight variations between samples in the amount of accidental annealing during soldering of electrical leads affects this comparison. It is, therefore, possible to derive only a rather rough value of the activation energy:

$$E=0.6 \text{ ev}.$$

A considerable fraction of the radiation effect is still retained at these temperatures and first begins to anneal out appreciably in the temperature range 300–350°C. Annealing curves for 300°, 325°, and 350°C are shown

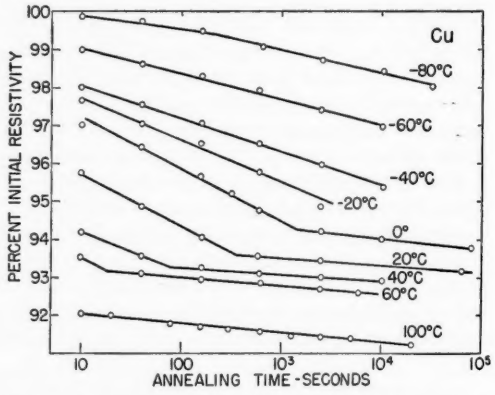


FIG. 5. Annealing of resistivity induced in 0.013-in. Cu wires by neutron irradiation. Measurements at 80°K. Curves indicate annealing processes with approximately logarithmic decrease of resistance with time.

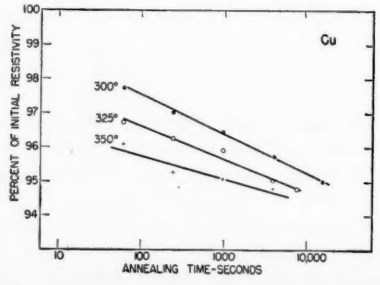


FIG. 6. Annealing of resistivity induced in 0.013-in. Cu wires by neutron irradiation. Initial resistivity refers to the state after long anneal at 20°C to complete 11 percent reduction in resistivity by the low-temperature process.

in Fig. 6. Initial resistivity in this case refers to the state of samples irradiated at low temperature but subsequently stored for 48 hours at room temperature to allow the lower-temperature process to go substantially to completion. Since the shortest annealing times at high temperature were one minute, only the latter parts of the curves appear in Fig. 6. It is nevertheless possible to derive an approximate value for the activation energy of this higher-temperature process:

$$E = 2.0 \text{ ev.}$$

Critical Shear Stress

Critical shear stress was measured from stress-strain curves, or rather torque *vs* deflection curves for the torsion pendulum. Typical data are shown in Fig. 7, where the horizontal scale represents maximum shear strain at the surface where the first plastic flow occurs. Irradiation is seen to produce an increase of approximately a factor of 6 in the critical shear stress of Al. As compared with resistivity, critical shear stress is an inherently imprecise quantity, depending on the extent of deviation from elastic behavior chosen as a criterion of beginning of plastic flow. Using the criterion of 2×10^{-5} plastic strain, a series of stress-strain curves were taken after different annealing times. In Fig. 8 each point represents the critical shear stress determined from one stress strain curve. As in the case of resistivity data, a heat of activation can be determined by plotting $\ln t$ *vs* $1/T$ for equal amounts of annealing at the three temperatures. These points, shown in Fig. 4, demonstrate that annealing of both resistivity and critical shear stress is governed by the same heat of activation, since the families of lines lie parallel to each other. Comparison of the two sets of curves shows further that longer times are required to reach a given stage of annealing in critical shear stress. It may be concluded that a smaller number of residual defects are necessary for the same relative effect, that is, the radiation effect on critical shear stress approaches saturation more nearly than for resistivity.

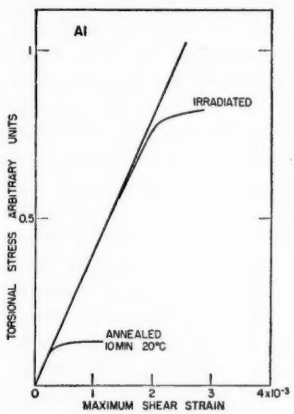


FIG. 7(a). Torsional stress *vs* strain for Al wires after 30-day neutron irradiation at 80°K and after subsequent 10-minute room temperature anneal. Strain refers to the maximum value at wire surface.

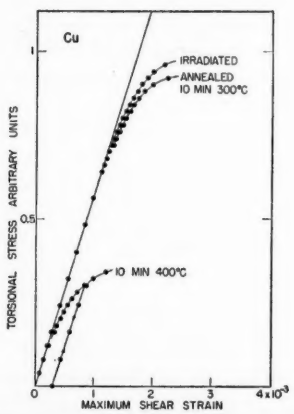


FIG. 7(b). Torsional stress *vs* strain for Cu wires after 30-day irradiation at 80°K, and after high-temperature anneals. Critical shear stress values were determined from each curve.

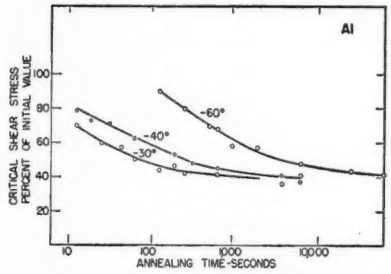


FIG. 8. Annealing of Al wires hardened by 30-day neutron irradiation at 80°K. Each critical shear stress point is determined from the torsional stress-strain curve on one sample, as in Fig. 7(a).

The annealing effects on critical shear stress in copper were similarly measured, with results as shown in Figs. 7(b), 9, and 10. As for aluminum a large increase in strength results from irradiation, but in copper very little annealing occurs below 300°C [Fig. 7(b)]. In Fig. 9 each point is derived from one stress-strain curve, the samples having been annealed 10 minutes each at different temperatures. The shear-strength drops rapidly between 300° and 400°C then more slowly. Further softening at 500°C and above cannot be unambiguously attributed to removal of lattice defects since at these temperatures some grain growth is also possible. A series of critical shear stresses were determined after different times and temperatures of annealing as seen in Fig. 10. The isothermal annealing curves are of a general form consistent with decay of irradiation effects according to either Eq. (2) or Eq. (3). As for aluminum the annealing of resistivity occurs somewhat more rapidly but nearly enough in the same temperature range to establish that both result from the same process. The activation energy determined as

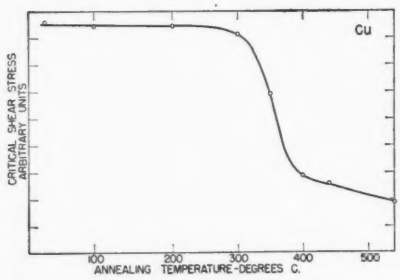


FIG. 9. Annealing of Cu wires hardened by 30-day neutron irradiation at 80°K. Each sample was annealed 10 minutes at the indicated temperature.

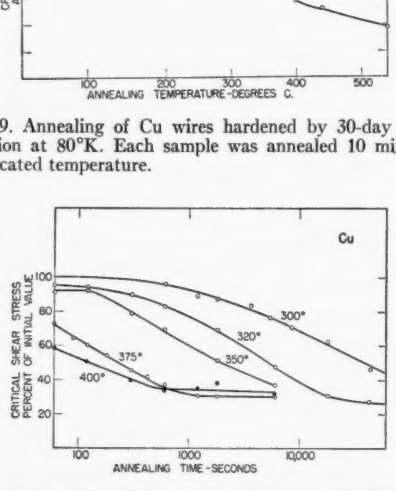


FIG. 10. Annealing of Cu wires hardened by 30-day neutron irradiation at 80°K. Each point was determined from one stress-strain curve, as in Fig. 7(b).

described above is $E=1.9$ ev. The activation energies for the various annealing properties studied may be summarized in Table I.

Other Physical Properties

The irradiation effects on some other physical properties were investigated to the extent of determining their suitability as an index of radiation damage. In each case the effect was found not suitably large compared to expected accuracy of measurement. The torsional rigidity modulus was measured by the natural frequency of the torsional pendulum and found not to change on annealing by an amount greater than the $\frac{1}{2}$-percent reproducibility attainable in replacing samples after annealing. Internal damping effect likewise was found not to exceed the somewhat larger experimental error resulting from extreme sensitivity of such samples to slight plastic deformation. Warren[7] found that Cu–Si single crystals irradiated at the same time as the Cu and Al specimens and examined subsequently at Massachusetts Institute of Technology showed negligible x-ray diffraction line broadening.

Discussion

The observed annealing effects in copper occur in two more or less distinguishable stages up to room temperatures, with a third stage at around 300°C. At the lowest temperatures studied, −80° to −40°C, the time rate of decrease of resistivity is proportional to $1/t$ over the entire range, beginning at a few seconds. Thus the rate is initially very high but rapidly decreases. Such behavior could be accounted for as a distribution of activation energies, ranging from 0 to 0.6 ev, the lower energy processes going to completion in the early stages. Alternatively, with the more specific model of vacancy-interstitial recombination, it can be accounted for by the additional assumption of a distribution of initial spacing between pairs. Annihilation would become slower as close spaced pairs disappeared first. In the next temperature range, −20° to 40°C the curves conform more nearly to a bimolecular form, [Eq. (3)] with a single activation energy of around 0.6 ev. These results seem consistent with several recent studies on resistivity effects from deuteron bombardment by Overhauser[2] and by Marx,[3] from α-particle bombardment by Eggleston,[4] and from mechanical deformation by Manintveld[8] and by Eggleston.[9] These various investigations, which are reviewed by Broom,[10] all showed a low activation energy of 0.2–0.25 ev at temperatures ∼−100°C, then a second activation energy of 0.67–0.88 ev at temperatures ∼0°C. Overhauser[2] found activation energies increasing linearly with temperature up to −30°C above which a single energy prevailed. His one isothermal curve, at −18°C corresponded to Eq. (1) with $n=2.5$ rather than $n=2$ as for a bimolecular recombination. A distribution in initial pair spacings, which he suggests to explain the apparent spread of activation energies, might also be expected to distort the shape of the isothermal annealing curve, however.

Those lattice defects which are annealed out in the above low-temperature stages are responsible for the major part of the radiation induced excess resistivity. It is interesting, however, that their removal has negligible effect on the critical shear stress. It remains high until, around 300°C, there is reached another annealing stage in which both the remaining excess resistivity and the increase in critical shear stress are removed, presumably by disappearance of a different type lattice defect. The associated activation energy of ∼2 ev is in accord with Eggleston's observed 2.12-ev energy for α-particle induced resistivity effect, as well as Redman, Coltman, and Blewitt's[11] measurement of a

TABLE I. Activation energies for annealing of neutron irradiation effects.

Metal	Temperature range °C	Activation energy—ev Resistivity	Critical shear stress
Al	−80° to −20°	0.55	0.55
Cu	−80° to −20°	0.6	(no effect)
	300° to 350°	2.0	1.9

[7] B. E. Warren (private communication).

[8] J. A. Manintveld, Nature **169**, 623 (1952).
[9] R. R. Eggleston, J. Appl. Phys. **23**, 1400 (1952).
[10] T. Broom, Phil. Mag. **3**, 26 (1954).
[11] Redman, Coltman, and Blewitt, Phys. Rev. **91**, 448 (1953).

2.2-ev activation energy in critical shear stress effect of neutron irradiation on copper single crystals. All three values are in agreement with the energy 2.07 ev for self-diffusion in copper as reported by Meier and Nelson.[12]

In the case of aluminum, where the temperature ranges for annealing are lower, the situation is somewhat simpler than for copper, since only one stage of annealing is observed. Also, there is unfortunately little previous work for comparison. By analogy with copper, also a face-centered structure, one might expect a qualitatively similar sequence of annealing stages. For aluminum, both resistivity and critical shear stress are restored to the fully annealed state in the $-80°$ to $-40°$C process, which would seem therefore to be comparable to the 300°–350°C process in copper. There is further evidence of the correspondence of these two stages in that the aluminum isothermal annealing curve follows the bimolecular form of Eq. (3), consistent with the high temperature data for copper but in definite contrast to curves for copper at room temperature and below. A third test of the analogy between copper and aluminum would be equality of the activation energies for the annealing process and for self-diffusion, but data for aluminum self-diffusion are lacking, primarily because of the absence of a suitable radioactive tracer isotope. Indirect values from magnetic resonance methods by Seymour[13] and from consideration of dilute solution diffusion by Nowick[14] give 0.9 ev and 1.4 ev, respectively. Both are much greater than the 0.55 ev found for annealing of radiation effects. If the analogy is nevertheless considered valid, other stages of radiation annealing in aluminum are to be expected at much lower temperatures. These would not have been observed in our experiments with $-150°$C irradiation.

Although our data plus those of the various references quoted show a consistent pattern in the annealing kinetics of resistivity and critical shear stress of the two materials, they are still hardly sufficient for definite identification of the lattice defects involved. One may set up a rather speculative general model however. The initial state of the irradiated metal must involve a distribution of interstitials and vacancies. With increasing temperature one of these will become mobile enough to diffuse through the lattice. According to Huntington's[15] calculations the interstitials should have lower activation energy (around 0.25 ev) and diffuse first. Removal of defects can then proceed by annihilation of pairs. A spread of activation energies may be expected both from the initial nonuniform distribution and from the fact that both vacancy and interstitital diffusion is involved. These single defects have large effect on resistivity but not necessarily on shear strength. Not all of the single defects are removed by recombination however, many diffusing instead to dislocations and either by settling there or by changing the shape and position of the dislocation edge, impeding its motion. At this stage the configuration is relatively stable and results in some increase in resistivity and a large increase in shear strength. In order for further annealing to occur, temperatures must be reached at which appreciable self-diffusion occurs, that is, at which thermal energy is sufficient for both creation and diffusion of vacancies.

Observations and tentative interpretations are summarized below:

(1) Neutron irradiation with total flux 1.1×10^{19} n/cm^2 causes increases in electrical resistivity of copper and aluminum by about 20 and 33 percent, respectively, and increases by about a factor of 5 in critical shear stress for both.

(2) In the range $-80°$ to $+20°$C, two-thirds of the resistivity effect in copper anneals out with no decrease in the critical shear stress effect. This process is to be associated with the removal from the lattice of vacancy and interstitial defects by diffusion and probably annihilation. There is evidence of a spread of activation energies up to 0.6 ev.

(3) In the range 300° to 350°C another stage of annealing in copper removes the remaining resistivity effect and all of the critical shear stress effect simultaneously, with activation energy 1.9 to 2 ev. This process is associated with ordinary self-diffusion, where creation of new vacancies and their diffusion is made possible by thermal energy. The shear strength removed in this process probably results from impedance of motion of dislocations by their interaction with vacancies and interstitials which diffused to them in the lower temperature processes.

(4) In aluminum both the resistivity and critical shear stress effects of irradiation anneal out together in the temperature range around $-60°$C, with activation energy 0.55 ev. There is evidence that this process is analogous to the 320° process in copper. Other annealing stages are, therefore, to be expected at low temperatures.

The authors are indebted to R. A. Meyer for design work on the irradiation facility and to G. H. Vineyard, G. J. Dienes, and others for discussions of the interpretation of the results.

[12] M. S. Meier and H. R. Nelson, Trans. Am. Inst. Mining Met. Engrs. **147**, 39 (1942).
[13] F. W. Seymour, Proc. Phys. Soc. (London) **A66**, 85 (1953).
[14] A. S. Nowick, J. Appl. Phys. **22**, 1182 (1951).
[15] H. B. Huntington, Phys. Rev. **91**, 1092 (1953).

Paramagnetic Resonance Spectrum of Some Doubly Activated Phosphors

W. Low
Department of Physics, The Hebrew University, Jerusalem, Israel
(Received January 6, 1955)

The paramagnetic resonance spectra of four different infrared stimulable strontium sulfide phosphors have been investigated. The hyperfine spectrum in SrS:Eu,Sm phosphor confirms the existence of divalent europium. The quantum efficiency in this phosphor is less than 1 photon per 10 europium activators. The ground state of europium is split by 0.18 cm^{-1} because of spin-orbit coupling in a cubic field. The manganese spectrum in a solid solution of strontium sulfide and selenide is discussed.

INTRODUCTION

THE processes involved in fluorescence and phosphorescence of inorganic phosphors are at present imperfectly understood. Most of these phosphors contain paramagnetic activators, which form luminescent centers or traps. These paramagnetic ions are in a definite valence state and may lose or gain electrons in these processes. In many cases the valency of these activators is not known with certainty or can only be inferred with difficulty from chemical or spectroscopic evidence. Often there are auxiliary activators present in the phosphor. These may have been incorporated without knowledge or design into the lattice of the host material, and owe their origin to the impurity of the chemical substances of which the phosphor is made. Moreover, there is at present no direct way of measuring the number of activators taking part in these processes.

Paramagnetic resonance is an ideal method for studying the valence state of paramagnetic impurities, as each ion of different valence state has a separate spectrum. It gives, moreover, information as to the number of paramagnetic ions per unit cell and an indication of their location in a single crystal. Unfortunately it is very difficult to grow single-crystal phosphors. In general, phosphorescent substances are available only in powdered form.

As the paramagnetic spectrum is a function of the orientation of the crystalline electric field with respect to the external magnetic field, the resonance lines are very much broadened due to the random orientation of the crystallites. Only in a few cases, where the spectrum shows little anisotropy, is information easily obtained. This fact limits the impurities which can be studied in this manner to a restricted group, in the main to divalent manganese, divalent europium, and trivalent gadolinium. These paramagnetic ions have an S ground state; their spectra are centered around $g=2.00$ (g is the gyromagnetic ratio of the electron spin) and can be studied at room temperature.

This paper reports some new results on powdered phosphors which contain divalent europium or divalent manganese. These strontium sulfide phosphors are activated by two impurities, a dominant activator, presumably the luminescent center, and the auxiliary activator, probably the electron trap. The crystal structure of strontium sulfide is of cubic symmetry,[1] and, therefore, no undue complications caused by low-order symmetries of the crystalline field were anticipated in the interpretation of the spectrum.

The paramagnetic spectrum of divalent manganese has been intensively studied. The ground state is $^6S_{5/2}$. Experiments of Bleaney and Ingram[2] on manganese salts and the interpretation of the spectrum by Abragam and Pryce[3] have shown that the six-fold degeneracy is removed mainly because of the noncubic portion of the field in the salts studied. A very pronounced hyperfine structure due to the magnetic moment of Mn^{55} has been found. This has been explained as arising from an unpaired electron because of configurational interaction in which a $3s$ electron is promoted to the $4s$ state.

Schneider and England[4] and Hershberger and Leifer[5] have investigated ZnS phosphors activated with manganese. They find that hexagonal ZnS gives a complicated spectrum, while cubic ZnS shows only the central six lines so characteristic of the hyperfine structure spectrum of manganese. In the phosphor discussed here, the crystal host consisted of a solid solution of cubic strontium sulfide and strontium selenide. It was interesting to find out to what extent the diamagnetic neighbors influence the hyperfine structure.

Divalent europium has a half-filled $4f$ shell and, therefore, a ground state of $^8S_{7/2}$. In a cubic field the ground state is split into a twofold degenerate level, followed by a fourfold, and then again by a twofold degenerate level. The ratio of the separation of the levels is 5:3. A theory of the splitting of the energy levels in the presence of an external field for various directions of the crystalline axes with respect to this field has been developed by DeBoer and van Lieshout[6]

[1] R. W. G. Wycoff, *Crystal Structures* (Interscience Publishers, New York, 1948).
[2] B. Bleaney and D. J. E. Ingram, Proc. Roy. Soc. (London) **A205**, 336 (1951).
[3] A. Abragam and M. H. L. Pryce, Proc. Roy. Soc. (London) **A205**, 135 (1951).
[4] E. E. Schneider and T. S. England, Physica **17**, 221 (1951).
[5] W. D. Hershberger and H. N. Leifer, Phys. Rev. **88**, 714 (1952).
[6] J. DeBoer and R. Van Lieshout, Physica **15**, 568, (1949).

for trivalent gadolinium in a cubic field. This theory should also apply to divalent europium. A hyperfine spectrum resulting from the magnetic moments of the two isotopes of europium was found by Bleaney and Low[7] while investigating one of these phosphors. This paper is not concerned with the nuclear information derived from this spectrum but mainly with the implications of the spectrum as far as these phosphors are concerned.

PROPERTIES OF THE PHOSPHORS INVESTIGATED

The phosphors for which the paramagnetic resonance spectra were examined are all infrared-stimulable. They all have two or more rare earth activators. They are SrS:Eu,Sm (Std. VI); SrS:SrSe:Eu,Sm (B 1); SrS:Ce,Sm (Std. VII); and one experimental phosphor similar to B 1 but containing in addition Mn (and a little gadolinium). The dominant activators (Eu, Ce · · ·) constituted less than one part per ten thousand per mole of strontium sulfide, and the auxiliary activator (Sm) about 5–10 times this amount.

Considerable experimental work has been done on the optical,[8–11] photoconductive,[12,13] and electroluminescent[14] properties of these phosphors. Their optical characteristics can best be explained by a concrete example of Std. VI. This phosphor can be excited by blue or ultraviolet light. It stores the energy, and on subsequent illumination by infrared radiation it emits visible light. The usual explanation of the mechanism of infrared stimulability is illustrated in Fig. 1. Blue light ($h\nu_1$) excites the luminescent center, which is the dominant activator, in our example divalent europium is ionized and loses an electron to the conduction band. The electron is captured by a low-lying trap, the auxiliary activator (trivalent samarium). The energy level of samarium is only 1.25 ev below the conduction band, and infrared light ($h\nu_2$) can, therefore, ionize the auxiliary activator. The electron is captured by a trivalent europium ion and emits visible light ($h\nu_3$).

These events can be written as chemical reactions as follows:

$$\begin{aligned} &\text{Excitation:} && \mathrm{Eu^{2+}}+h\nu_1\rightarrow\mathrm{Eu^{3+}}+e^-, \\ &&& \mathrm{Sm^{3+}}+e^-\rightarrow\mathrm{Sm^{2+}}; \\ &\text{Stimulation:} && \mathrm{Sm^{2+}}+h\nu_2\rightarrow\mathrm{Sm^{3+}}+e^-, \\ &&& \mathrm{Eu^{3+}}+e^-\rightarrow\mathrm{Eu^{2+}}+h\nu_3. \end{aligned} \qquad (1)$$

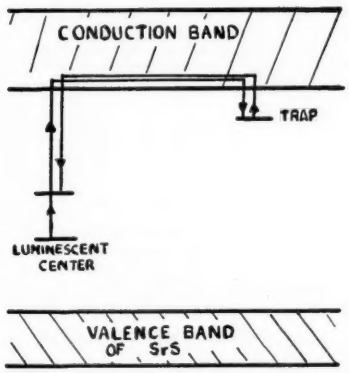


FIG. 1. Schematic picture of energy bands and levels in an infrared storage phosphor.

In the case of Std. VII the cerium changes valencies from a trivalent to a quadrivalent state.

This picture accounts for the fact that the wavelength of the emitted light ($h\nu_3$) differs for the various phosphors, whereas the peak stimulation frequency is substantially the same for all these phosphors. The emitted light is a characteristic of the dominant activator, whereas the stimulation frequency is a function of the auxiliary activator Sm, common to all these phosphors.

To substantiate this picture it would be necessary to show that the activators are in the unexcited state in their proper valencies and that these valencies change on excitation. There is some evidence[9] that there is trivalent samarium in the unexcited phosphor. The evidence that europium is in the divalent state is not so conclusive. No characteristic line spectrum,[15] so peculiar to trivalent europium, has been found.

EXPERIMENTAL RESULTS

All measurements were made at 3 cm, and a range of 0–6000 gauss was scanned, at room, liquid oxygen, and liquid hydrogen temperatures. The apparatus was the standard 3-cm setup used at the Clarendon Laboratory.[2] In the case of Std. VI, the influence of ultraviolet light on the paramagnetic spectrum was investigated. This was done by removing the sample from the cavity, placing it in a wide glass vessel, and exposing it to very strong ultraviolet light for half an hour. The glass vessel was shaken off and on to make sure that as many grains as possible were exposed to the radiation. The signal to noise ratio was compared before and after excitation. While the results were not entirely reproducible, because of instability of the oscillator, and because the cavity was not always filled to the same level, one can conservatively estimate that changes of 10 percent would have been observable.

Std. VI.—Three phosphors of slightly different europium-to-samarium ratio and of different phosphor

[7] B. Bleaney and W. Low, Proc. Phys. Soc. (London) **68**, 55 (1955).
[8] Urbach, Pearlman, and Hemmendiger, J. Opt. Soc. Am. **36**, 373 (1946).
[9] G. Fonda and F. Seitz, *Preparations and Characteristics of Solid Luminescent Materials* (John Wiley and Sons Inc., New York, 1948).
[10] P. Brauer, Z. Naturforsch. **1**, 70 (1946); **2a**, 238 (1947).
[11] F. Stockmann, Naturwiss. **10**, 226 (1952).
[12] Low, Braun, and Steinberger, J. Opt. Soc. Am. **44**, 88 (1954).
[13] J. J. Dropkin, "Photoconduction in Phosphors" Polytechnic Institute of Brooklyn, 1954 (unpublished).
[14] Low, Steinberger, and Braun, J. Opt. Soc. Am. **44**, 504 (1954). (A full report is in preparation.)
[15] R. Tomaschek, Ergeb. exakt. Naturw. **20**, 268 (1942).

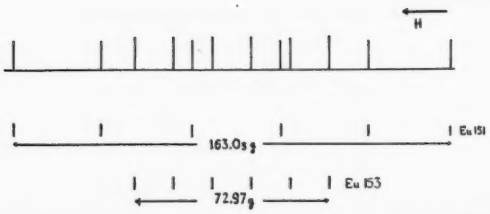


FIG. 2. The paramagnetic resonance spectrum of SrS:Eu,Sm. The hyperfine spectrum consists of two groups of 6 lines due to the two europium isotopes with nuclear spins of 5/2. It confirms the presence of divalent europium in this phosphor.

grain sizes were investigated. The same spectrum was obtained at all times and at all temperatures. The observed spectrum is shown schematically in Fig. 2. These twelve lines were superimposed on a very broad line of roughly 100 gauss total width. The spectrum was centered around $g=2.00$. The signal increased with the amount of sample in the cavity without any noticeable broadening. The lines were very narrow, the half-width at half-intensity being slightly less than 2 gauss. As will be discussed later, the spectrum can be resolved into two groups of six lines each of nearly equal intensity, corresponding to the relative abundance of the two europium isotopes (47.8 percent and 52.2 percent).[16]

B 1.—This phosphor, containing Eu and Sm in a solid solution of SrS and SrSe, shows a broad line of roughly 100 gauss centered around $g=2.00$. There seemed to be some components superimposed on this line but of such small intensity that no quantitative measurements were made. Only one sample was tested.

B 1 with additional manganese.—The spectrum is shown in Fig. 3, as seen at room temperature. At liquid hydrogen temperature the spectrum saturates and broadens so that only 6 lines are seen. The spectrum is centered around $g=2.00$. The half-width at half-intensity was about 4–5 gauss. The two center lines were broader and could not be resolved into two components. The spectrum can be reconstructed as two groups of 6 lines, with the center of gravity of these lines nearly coinciding. The separation of the outermost components is 475 gauss, corresponding to a hyperfine structure constant $A=95$ gauss; and of the next inner components 460 gauss, corresponding to $A=92$ gauss.

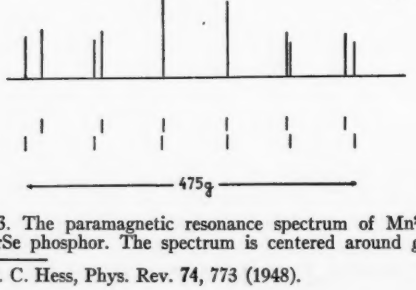


FIG. 3. The paramagnetic resonance spectrum of Mn^{2+} in a SrS:SrSe phosphor. The spectrum is centered around $g=2.00$.

[16] O. C. Hess, Phys. Rev. **74**, 773 (1948).

Std. VII.—This line shows a fairly strong line close to $g=2.00$ at room temperature (see Fig. 4). At liquid oxygen and liquid hydrogen temperatures, one can resolve two additional lines of smaller intensity asymmetrically placed with respect to the center line. Besides, at least two weaker components could be seen on the oscilloscope. In addition, there is at liquid hydrogen temperature a very intense, bell-shaped line of about 100 gauss total width at $g=1.33$. This line disappears at liquid oxygen temperature. Various phosphors were investigated, and in all phosphors essentially the same spectrum was obtained. Those phosphors which were made with LiF flux showed slightly greater absorption.

INTERPRETATION OF RESULTS

There are two kinds of spectra to be considered, those which contain only rare earth activators and those which have manganese in addition. In the former the magnetic carriers are the $4f$ electrons, and in the latter those in the $3d$ shell. In the ion group the $3d$

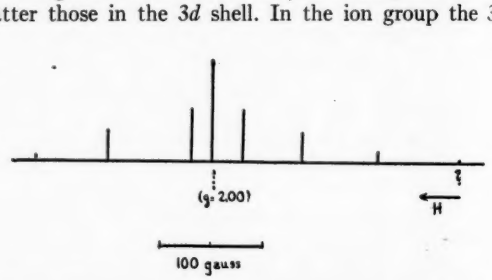


FIG. 4. The paramagnetic resonance spectrum of SrS:Ce,Sm. The spectrum is centered around $g=2.00$. The question mark indicates that the line was about of noise level intensity and could not be determined with any certainty.

shell is the outermost shell and is, therefore, exposed to very strong crystalline electric fields. The orbital component contributes only slightly to the spectroscopic splitting factor through spin-orbit coupling. The g factors of the iron group are, in general, close to two and show little anisotropy.

In the rare earth group the crystal field is much smaller than in the iron group and even less than the spin-orbit interaction. This is due to the fact that the $4f$ electrons are partially shielded by the $5s^25p^6$ electron shells and, therefore, do not feel the total neighboring charge, and is also due to the larger distance from the diamagnetic neighbors. The g factors in the rare earth group show considerable anisotropy and vary considerably from the value two. The relaxation time is, in general, much shorter than in the iron group and, therefore, in many cases can only be detected at very low temperatures.

Std VI (SrS:Eu,Sm).—As previously outlined, we may expect that the activators will have the following valencies: Eu^{2+}, Eu^{3+}, Sm^{2+}, Sm^{3+}. The ground state of Eu^{3+} and Sm^{2+} is $^7F_0$ and would not show any

resonance spectrum. The spectrum of Sm^{3+} has been found in an ethyl sulfate crystal.[17] It can only be detected at very high magnetic fields ($g_{\parallel}=0.596$, $g_{\perp}=0.604$) and at temperatures lower than liquid hydrogen. We would, therefore, not expect to detect this spectrum with the magnetic field and temperatures employed. The ground state of divalent europium is $^8S_{7/2}$. Its eightfold degeneracy is partially removed by means of the spin-orbit coupling in the cubic field. DeBoer and van Lieshout[6] calculated theoretically the energy levels in an external magnetic field. They find that the spectrum is strongly dependent on the orientation of the magnetic field with respect to the crystalline electric field axes. In the powder only the $\frac{1}{2}\rightarrow-\frac{1}{2}$ transition is observed. The energy levels in a strong field are given by

$$E_{\pm\frac{1}{2}}=\pm\beta H+[(5/4)+(45/4)\phi]\delta\pm\alpha_1\delta^2/\beta H+\cdots,$$
$$\alpha_1=35/128-(1135/64)\phi+(10875/128)\phi^2-(3675/32)\psi, \quad (2)$$

where ϕ and ψ are the direction cosines of the magnetic field with respect the cubic axes x, y, z. That is,

$$\phi=a_x^2a_y^2+a_y^2a_z^2+a_z^2a_x^2,$$
$$\psi=a_x^2a_y^2a_z^2,$$

and 8δ is the over-all splitting at zero magnetic field due to spin-orbit interaction in a cubic field. The central transition then is given by

$$h\nu=E_{+\frac{1}{2}}-E_{-\frac{1}{2}}=g\beta H+2\alpha_1\delta^2/\beta H, \quad (3)$$

where $g=2.0023$.

The spectrum observed confirms the presence of Eu^{++}. The narrow lines indicate that the spectrum must indeed be isotropic. The measured g factor is $1.991_2\pm0.001$, very close to 2.0023. Assuming that the deviation from the free spin value is entirely due to spin-orbit coupling (and not to spin-spin interaction arising in a distorted electron distribution), we can calculate the zero-field splitting. Using $\bar{\phi}=\frac{1}{2}$, $\langle\phi^2\rangle_{Av}=1/21$, $\bar{\psi}=1/1.5$, one finds $\delta=0.023$ cm^{-1} and the over-all splitting $8\delta=0.18$ cm^{-1}. There are no experimental data from susceptibility measurements which might give the zero-field splitting for divalent europium. In trivalent gadolinium, which has the same ground state as divalent europium, in the salt $Gd_2(SO_4)_3\cdot8H_2O$, Benzie and Cooke[18] find a splitting of 0.94 cm^{-1}. In the paramagnetic spectrum of gadolinium ethylsulfate, Bleaney *et al.*[19] find that the main splitting is due to spin-spin interactions in the noncubic field of these crystals. Nevertheless, they find that the g factor is also 1.991, indicating that the deviations of the g factor in both substances is probably due to spin-orbit coupling. This splitting should be compared with the cubic contribution of 0.0134 cm^{-1} in the ferric alums[20] and the even smaller splitting of 0.001 cm^{-1} in the manganese salts.[2] This seems to support the contention that the splitting observed here is mainly due to spin-orbit coupling which is much weaker in the iron group.

The nuclear information obtained from the hyperfine structure has been discussed elsewhere.[7] The spectrum can be described by the formula:

$$h\nu=g\beta H+Am+\tfrac{1}{2}(A^2/h\nu)\{I(I+1)-m^2\}, \quad (4)$$

with $g=1.991_2\pm0.001$, $A_{151}=30.8_2\pm0.2$, $A_{153}=13.7_9\pm0.2$ (A in units of 10^{-4} cm^{-1}). It confirms[21] the nuclear spin of 5/2 for the two isotopes and gives the anomalous large ratio of the nuclear magnetic moments as $\mu_{151}/\mu_{153}=2.23_5\pm0.03$. The hyperfine structure probably has its origin in the configurational coupling of $4s^2\,4f^7\,5s^2$ and $4s^2\,4f^7\,5s\,6s$. This will not appreciably influence the theory of Elliott and Stevens.[22] Rough calculations on erbium and other rare earth ethylsulfates, taking into account a configurational interaction of this magnitude, show a slightly better agreement between the theory and the experimental data.

The narrow lines observed with SrS might make it possible to observe the spectrum of trivalent gadolinium and see whether any hyperfine structure can be resolved. It is planned to try to make such a phosphor.

The quantum efficiency of some of these phosphors has been studied. Urbach[23] reports that in some of Std. VI phosphors there is one photon emitted per three samarium activators. Ellickson,[24] studying Std. VII, comes to the conclusion that only 1 photon is emitted per 8 samarium ions, and in his phosphor this corresponds to 1 photon per 60 cerium activators. In all these phosphors the dominant activator was present in larger quantities than the auxiliary activator. Their measurements involve the assumptions that (a) in the preparation of the phosphors the mole ratio of the components has not changed, i.e., that all the activators have been incorporated in the strontium sulfide lattice; and (b) that all the activators have the proper valencies before excitation so that they can contribute to the phosphorescence. If either of these assumptions should not be correct, the efficiency would be higher.

The failure to observe any change in the signal-to-noise ratio in our measurements seems to indicate that the efficiency is less than one photon per 10 europium ions. Unlike Urbach's and Ellickson's phosphors, these phosphors have a larger amount of samarium than europium as activators.

[17] G. S. Bogle and H. E. M. Scovil, Proc. Phys. Soc. (London) **A65**, 368 (1952).
[18] R. J. Benzie and A. H. Cooke, Proc. Phys. Soc. (London) **A63**, 213 (1950).
[19] Bleaney, Scovil, and Trenam, Proc. Roy. Soc. (London) **A223**, 15 (1954).
[20] B. Bleaney and R. S. Trenam, Proc. Roy. Soc. (London) **A223**, 1 (1954).
[21] H. Schuler and T. Schmidt, Z. Physik **94**, 547 (1935).
[22] R. J. Elliott and K. W. H. Stevens, Proc. Roy. Soc. (London) **A218**, 553 (1953).
[23] See reference 9, pp. 119.
[24] R. T. Ellickson, J. Opt. Soc. Soc. Am. **36**, 501 (1946).

B 1. (SrS:SrSe,Eu,Sm).—No information can be obtained. The spectrum is probably similar to Std. VI.

B 1 with manganese (SrS:SrSe,Eu,Sm,Mn).—The spectrum of Mn^{2+} has been discussed in detail by Bleaney and Ingram.[2] Abragam and Pryce[3] showed that the spectrum in a pure cubic field could be represented by

$$\mathcal{H}=g\beta\mathbf{H}\cdot\mathbf{S}+(a/6)(S_1^4+S_2^4+S_3^4)+A(S_xI_x)+B(S_xI_x+S_yI_y), \quad (5)$$

where g is the spectroscopic splitting factor, β the Bohr magneton, H the external field, S the electron spin operator, I the nuclear spin operator, a the measure of zero-field splitting resulting from the cubic field interaction which has terms proportional to S_1^4, S_2^4, S_3^4, where the S_i are the components referred to the cubic axes. A and B are hyperfine structure constants of nearly equal magnitude. In case there is a noncubic portion of the symmetry, the energy levels will be split much more as the operator in this case would be quadratic in S.

The hyperfine structure of Mn^{55} has nearly a constant value in various salts investigated and is isotropic within a few percent.

With a powder of cubic structure only the $\frac{1}{2}\rightarrow-\frac{1}{2}$ transition is observed,[5,25] with its characteristic 6 hyperfine lines.

The spectrum observed, as shown in Fig. 3, shows in the main the spectrum of Mn^{2+} incorporated in a cubic lattice of SrS. The splitting of the 6 lines into narrow doublets, with the center of gravity nearly constant, is somewhat puzzling. According to Abragam,[26] the hyperfine structure is due to configurational interaction and should, therefore, be the property of the free ion. Possibly these two groups of 6 lines are due to the two nonequivalent manganese ions, one surrounded by sulfur, the other by selenide neighbors. This would indicate that the amount of configurational interaction might be dependent to a small degree on the crystalline surrounding. If a distortion in the cubic field be present, this would also produce a hyperfine structure but probably less than 10^{-4} cm^{-1}, or less than 1 gauss.

The sulfide to selenide ratio in these phosphors was of the order of 1:5. The signals of the two groups of lines was, however, of nearly equal intensity. If the above interpretation is correct, this would indicate that Mn^{2+} is more easily substituted in the strontium sulfide than in the strontium selenide. Further experiments with widely different ratios of selenide to sulfide might clarify this point.

Std. VII (SrS:Ce,Sm).—Cerium can be in the divalent, trivalent, and quadrivalent states. In the quadrivalent state it is diamagnetic. In the trivalent state the ground state is $^2F_{5/2}$ and in the divalent state if it exists the ground state is $^3H_4$ (as in trivalent praseodymium). No measurements on a single crystal of cubic symmetry have been made so far. In trivalent cerium, in a single crystal of cerium ethylsulfate, Bogle *et al.*[27] find at liquid helium temperatures two resonance lines, one due to the $\pm5/2$ levels and the other to the $\pm\frac{1}{2}$ level, the latter lying lower. Their values are $g_{\parallel}(\frac{1}{2})=0.955$, $g_{\perp}(\frac{1}{2})=2.183$, $g_{\parallel}(5/2)=3.72$, $g_{\perp}(5/2)=0.2$. Lancaster and Gordy[28] find in powdered concentrated cerium oxalate a spectrum at $g=1.95$ at room temperature with the narrow line width of 170 gauss. Similarly, Kurenev and Salikhov[29] report a spectrum at room temperature in cerium sulfate. Cerium has no odd isotopes and, therefore, no hyperfine spectrum.

We cannot account for the spectrum shown in Fig. 4 with certainty. It is possible that the central strong line with g slightly less than 2 is due to the trivalent cerium level of $\frac{1}{2}$. This would be very similar to the line observed by Lancaster and Gordy. As the g factor seems to be very anisotropic, the $g_{\perp}$ would be very strongly weighted corresponding to the fact that the majority of ions lie perpendicular to the field. This would peak the line near $g=2.00$. But as rare earths are difficult to obtain with very high purity, one cannot be sure whether the central line may not be the results of gadolinium or possibly neodymium ($g_{\perp}=2.073$), which may be present as small impurities. The other weak lines are probably impurities as well.

We similarly cannot assign the very strong line at $g=1.33$ seen only at liquid hydrogen temperatures. It is possible that this line may be due to praseodymium, which has been shown to have a hyperfine structure of about 90 gauss.[30,31] This illustrates how difficult it is to get any definite information from powdered specimens.

CONCLUSIONS

The work with strontium sulfide phosphors has indicated that quantitative information in powders can be obtained only with magnetic activators which have a very isotropic spectrum, like ions with an S ground state. With these ions one finds very narrow lines. In the case of europium-activated phosphors, we confirm the presence of divalent europium and could calculate the zero-field splitting. The quantum efficiency is less than 1 photon per ten europium activators. An apparatus is being set up which will make it possible to measure the changes on excitation (by x-rays) *in situ* and thus determine the efficiency with greater accuracy.

The narrow lines seen give hope that a hyperfine structure in gadolinium may be observed. We have

[25] J. S. Van Wieringen, Physica **19**, 397 (1953).
[26] A. Abragam, Phys. Rev. **79**, 534 (1950).
[27] Reported by B. Bleaney and K. W. H. Stevens, Repts. Progr. Phys. **16**, 149 (1952).
[28] F. W. Lancaster and W. Gordy, J. Chem. Phys. **19**, 1181 (1951).
[29] V. Y. Kurenev and S. G. Salikhov, Zhur. Exsptl. Teort. i Fiz. **21**, 864 (1951).
[30] B. Bleaney and H. E. D. Scovil, Phil Mag. **43**, 999 (1952).
[31] Davis, Kip, and Malvano, Atti accad. nazl. Lincei **11**, 77 (1951).

embarked on a systematic study of all rare earth activators in strontium sulfide. An attempt will be made to grow single crystals.

Part of this work was done while the author spent the summer at the Clarendon Laboratory, Oxford.

The author should like to thank Dr. B. Bleaney for permission to use his paramagnetic resonance spectrometer and for many helpful discussions. The author is also indebted to Dr. L. Weidenfeld for furnishing the phosphors.

PHYSICAL REVIEW VOLUME 98, NUMBER 2 APRIL 15, 1955

Temperature Dependence of the Width of the Band Gap in Several Photoconductors

RICHARD H. BUBE
Radio Corporation of America, RCA Research Laboratories, Princeton, New Jersey
(Received December 17, 1954)

The temperature dependence of the width of the forbidden band gap has been determined, between 90° and 400°K, for cadmium sulfide, cadmium selenide, cadmium telluride, and zinc selenide, by measurement of photosensitivity as a function of wavelength. The location of the maximum photosensitivity, often very sharp, at the wavelength corresponding approximately to the width of the band gap, makes this method very convenient for the determination. Structure exists in the photosensitivity maxima of cadmium sulfide and cadmium selenide, suggesting the existence of double conduction or valence bands.

INTRODUCTION

THE temperature dependence of the width of the forbidden gap of a semiconductor is commonly determined by measuring the dependence of the absorption on wavelength, at a number of different temperatures. Such measurements have been made by Kroeger,[1] Seiwert,[2] and Hoehler[3] for cadmium sulfide. Their results indicate that the absorption-edge shifts to longer wavelengths at the rate of about 1 A/°K for increasing temperature.

Because of the broadness of the absorption edge in most semiconductors, it is somewhat difficult to determine a width for the forbidden gap accurately, and hence to determine accurately the temperature dependence of the gap width. In many cases, such striving for greater accuracy may be pointless, because of an inherent lack of definition in the band edges themselves. But, at least experimentally, it is much simpler to determine a quantity which is very nearly equal to the width of the forbidden gap for photoconducting crystals, by measuring the photosensitivity as a function of wavelength. Most photoconducting crystals show a rather sharp maximum of photosensitivity at a wavelength corresponding closely to the width of the forbidden gap, as determined from measurements of absorption. Niekisch[4] and Caspary and Mueser,[5] have determined the temperature dependence of the forbidden gap for cadmium sulfide from measurements of photosensitivity as a function of wavelength. Their results also indicate approximately a 1 A/°K shift in absorption edge.

The purpose of this paper is to present data on the variation of the width of the forbidden gap with temperature between 90° and 400°K, for crystals of cadmium sulfide, cadmium selenide, cadmium telluride, and zinc selenide.

EXPERIMENTAL

Materials

The preparation of single crystals of "pure" cadmium sulfide, cadmium selenide, and zinc selenide by S. M. Thomsen, by reaction between the elements in the vapor phase, has been described in a previous publication.[6] The growth of crystals of cadmium telluride by D. A. Jenny and E. B. Lawton, by slow cooling from the melt, has also been previously reported.[7]

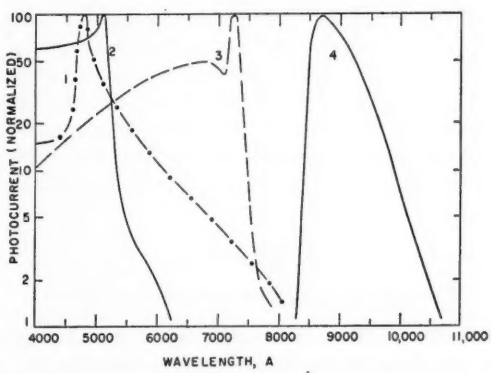


FIG. 1. Spectral curves of photosensitivity for (1) zinc selenide, (2) cadmium sulfide, (3) cadmium selenide, and (4) cadmium telluride, at room temperature.

[1] F. A. Kroeger, Physica **7**, 1 (1940).
[2] R. Seiwert, Ann. Physik **6**, 241 (1949).
[3] G. Hoehler, Ann. Physik **4**, 371 (1949).
[4] E. A. Niekisch, Ann. Physik **8**, 291 (1951).
[5] R. Caspary and H. Mueser, Z. Physik **134**, 101 (1952).
[6] R. H. Bube and S. M. Thomsen, J. Chem. Phys. **23**, 15 (1955).
[7] D. A. Jenny and R. H. Bube, Phys. Rev. **96**, 1190 (1954).

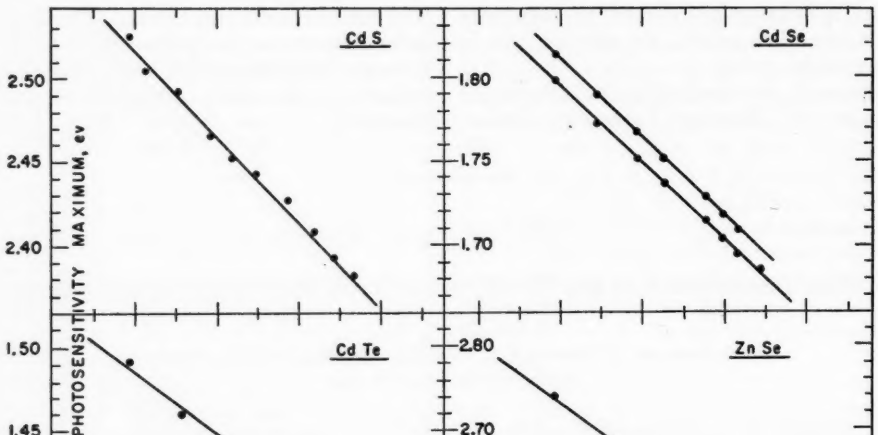


FIG. 2. The location of the photosensitivity maximum (or maxima) as a function of absolute temperature, for cadmium sulfide, cadmium selenide, cadmium telluride, and zinc selenide.

Electrical contact to the cadmium sulfide, cadmium selenide, and zinc selenide crystals was made using melted indium electrodes.[8] The ends of the crystal of cadmium telluride were plated with nickel and copper, and then leads were soldered to the plated portions.

Measurements

Photosensitivity spectra were measured with a 500-mm Bausch and Lomb grating monochromator, using the photoconductivity *vs* temperature apparatus previously described.[9] The highest resolution compatible with reliable measurements was used at all times; a resolution of 16.5 A was used for all measurements on cadmium sulfide and low-temperature measurements on cadmium selenide, 33 A for high-temperature measurements on cadmium selenide and all measurements on cadmium telluride, and 49.5 A for all measurements on zinc selenide.

RESULTS

Temperature Dependence of Gap Width

Spectral curves of photosensitivity at room temperature are given for the four different crystals in Fig. 1. In plotting the curves, the data were corrected so as to indicate the photocurrent for an equal number of photons incident per second at any wavelength, assuming a linear current-light relationship.[10] All of the crystals show a sharp photosensitivity maximum except cadmium telluride, and even for this crystal the identification of the maximum is not difficult.

The decrease in photosensitivity for wavelengths shorter than the wavelength for maximum photosensitivity for cadmium sulfide, cadmium selenide, and zinc selenide, is probably due to the fact that radiation in this range is heavily absorbed in surface regions which have an intrinsically smaller photosensitivity than volume regions. In the case of the cadmium telluride crystal, it is probable that the very rapid decrease of photosensitivity for wavelengths shorter than the absorption edge is caused both by this effect, and also by a geometrical effect whereby the applied field is concentrated across a section of the crystal somewhat distant from the surface upon which the radiation is incident.

The location of the photosensitivity maximum is plotted in energy units in Fig. 2 for each of the four crystals, as a function of temperature. The constants involved in a linear relationship between band gap in electron volts and absolute temperature were derived

[8] R. W. Smith and A. Rose, Phys. Rev. **92**, 857(A) (1953).
[9] R. H. Bube, J. Chem. Phys. **23**, 18 (1955).

[10] The assumption of a linear current-light curve is not justified for the cadmium selenide crystal, which has, for the light intensities used, a superlinear current-light dependence [see: R. H. Bube, paper in *Proceedings of the Conference on Photoconductivity* (John Wiley and Sons, New York, 1955)]. For this reason, the photosensitivity at longer wavelengths may be somewhat greater relative to that at shorter wavelengths than the curve in Fig. 1 for CdSe indicates.

from each of these plots by means of the method of least squares. Over this temperature range, an approximately linear relationship may also be derived between the wavelength corresponding to the absorption edge and the absolute temperature. These relationships are summarized in Table I.

Structure in the Photosensitivity Maximum

A detailed inspection of the photosensitivity maximum of cadmium sulfide, and of that of cadmium selenide (the two crystals for which the sharpness of the maximum and the resolution which could be used make such inspection meaningful), reveals that the photosensitivity maximum for both cadmium sulfide and cadmium selenide exhibits two components. The existence of such structure in the case of cadmium sulfide was suggested in the data previously reported by Klick.[11]

In Fig. 3(a)-3(e), the photosensitivity spectral response *vs* wavelength curve for cadmium sulfide is plotted in the wavelength range of the maximum for several temperatures. At the lowest temperatures, there is only a slight indication of structure in the curve, but at 218°K and above, the presence of two components,

TABLE I. Dependence of band gap (photosensitivity maximum) on temperature between 90° and 400°K.

Cadmium sulfide			
$E=2.57-0.00052T$ ev	and	$E=2.59-0.00052T$ ev	
$\lambda=4807+1.10T$ A	and	$\lambda=4772+1.10T$ A	
Cadmium selenide			
$E=1.84-0.00046T$ ev	and	$E=1.85-0.00046T$ ev	
$\lambda=6726+1.84T$ A	and	$\lambda=6661+1.84T$ A	
Cadmium telluride			
$E=1.52-0.00036T$ ev			
$\lambda=8110+2.23T$ A			
Zinc selenide			
$E=2.80-0.00072T$ ev			
$\lambda=4401+1.30T$ A			

separated by about 35 A is clear. The separation between the two components, which amounts to about 0.017 ev, does not change detectably between 218° and 341°K.

Similar results for cadmium selenide given in Fig. 3(f)-3(j), show that structure in the curve is detectable over the whole range of temperatures used. The separation between the components is about 65 A, apparently independent of temperature. In energy units, the separation is about 0.015 ev, very nearly

[11] C. C. Klick, Phys. Rev. **89**, 274 (1953).

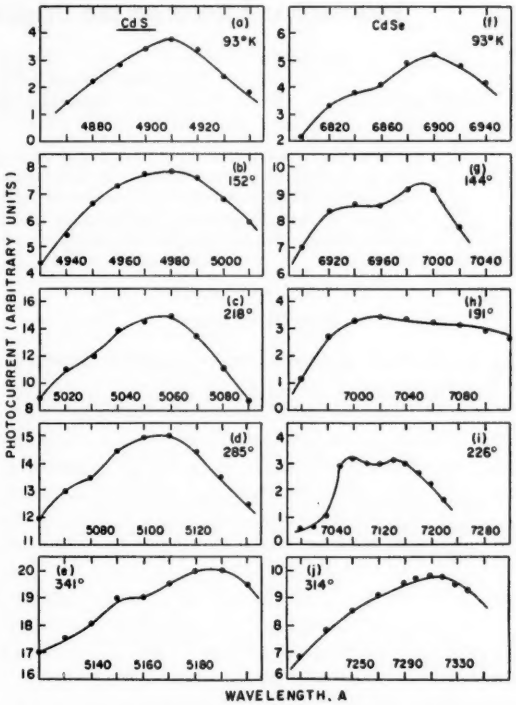


FIG. 3. The photocurrent (in arbitrary units) at the maximum of the photosensitivity *vs* wavelength curve for (a)–(e) cadmium sulfide, and (f)–(j) cadmium selenide, at several different temperatures.

the same as that found for cadmium sulfide. Both maxima have been plotted in Fig. 2 for cadmium selenide.

The relative magnitude of the two components for cadmium sulfide appears to be approximately independent of temperature. But for cadmium selenide, the low wavelength component increases in magnitude relative to the high wavelength component between 93° and about 200°K, and then decreases for temperatures between 200° and 314°K.

The existence of such structure in the photosensitivity maximum suggests the presence of double conduction or valence bands. No simple model of double bands seems able to describe adequately the experimental results. In particular, the temperature dependence of the relative magnitude of the two components for cadmium selenide seems to require an explanation involving more than optical transitions only.

The author is indebted to R. H. Parmenter and F. Herman for discussions on the nature and properties of multiple bands.

Polarization Effects on the Brightness Waves of Electroluminescence

FRANK MATOSSI
U. S. Naval Ordnance Laboratory, White Oak, Maryland
(Received December 6, 1954)

Electroluminescence excited by sinusoidal fields yields two maxima per half-period of the field. The intensity ratio of these maxima depends on field frequency and on the superposition of ultraviolet-excited luminescence. These observations are shown to be related to the existence of polarization effects, which are treated theoretically by extending previous theories of polarization effects in semiconductors (Jaffé and others), taking into account periodic excitations of centers by the field.

INTRODUCTION

RECENTLY, several authors presented theoretical treatments of polarization effects in semiconductors.[1,2] These are based on a set of differential equations governing the kinetics of electron and hole distribution under the influence of excitation, recombination, and diffusion processes in the presence of an electric field.[3] This theory was successfully applied to the understanding of some capacitance effects. From observations of electroluminescence excited by nonsinusoidal fields, it could be inferred that polarization effects are important also in electroluminescence.[4] Similar conclusions were presented by Curie[5] in discussing the brightness waves and other features of sinusoidally excited electroluminescence.

It seems proper, therefore, to relate some pertinent results of electroluminescence experiments with the theory of the polarization effects worked out by the aforementioned authors.[1-3] This will be done for sinusoidal fields only. By introducing into the equations of this theory a term that takes into account the periodic excitation of electrons to the conduction band by the periodic electric field, it is possible to describe some characteristic properties of the brightness waves.

OBSERVATIONS

The experimental results discussed are:

(1) In many electroluminescent phosphors (polycrystalline powders embedded in a dielectric) two brightness maxima appear for every half-period of the field. One is essentially in phase with the field, the other is out of phase. The former is called a *B*-maximum; the latter, a *C*-maximum, in conformance with the designation of related peaks observed with nonsinusoidal fields.[4] The existence of the *C*-peaks, which always occur when the external field is diminishing, is assumed to be due to the recombination of polarization charges diffusing back into the interior from the surface where they had been held up by the field; but also the intensity and the shape of the *B*-peaks is influenced by the polarization charges since these affect the internal field responsible for the electroluminescence processes. This interpretation[4] for the peaks observed in nonsinusoidal fields is considered to be applicable also to the two maxima of the sinusoidal case.

(2) If an electroluminescent phosphor is excited simultaneously by ultraviolet radiation and an electric field that is strong enough to excite electroluminescence without the ultraviolet, the brightness waves show the *B*-peaks only. The *C*-peaks appear only after removal of the ultraviolet excitation; however, not at once but only after some time in the order of several minutes. Figure 1 shows the difference of a brightness wave with and without simultaneous ultraviolet excitation.

(3) The ratio of the intensities of the *C*-peaks to the *B*-peaks increases with frequency so that the *C*-peaks at 1000 cps become as large as the *B*-peaks.

The specific results mentioned under (2) and (3) were obtained[6] with a phosphor prepared in the same manner as the phosphors used in the green Sylvania electroluminescent panel.[7] Similar qualitative results were also obtained with some other phosphors.

THEORETICAL CONSIDERATIONS

The equations governing the charge distribution in semiconductors are given by[1]

$$\begin{aligned}
&\partial p/\partial t = k_1 n_c - k_2 np + D_p(\partial^2 p/\partial x^2) - \mu_p[\partial(pE)/\partial x],\\
&\partial n/\partial t = k_1 n_c - k_2 np + D_n(\partial^2 n/\partial x^2) + \mu_n[\partial(nE)/\partial x],\\
&\partial n_c/\partial t = -k_1 n_c + k_2 np,\\
&\partial E/\partial x = 4\pi e(p-n)/\epsilon.
\end{aligned} \tag{1}$$

These equations describe the time rate of change of the number (per unit volume) of electrons (n) and holes (p). They take into account excitation of luminescence centers ($k_1 n_c$, where n_c = number of unexcited centers); recombination ($k_2 np$); diffusion (D_p, D_n = diffusion coefficients for holes and electrons, respectively); motion under the influence of an electric field E obeying Poisson's equation (μ_p, μ_n = mobilities, ϵ = dielectric constant, e = electronic charge). Transi-

[1] J. R. MacDonald, Phys. Rev. **92**, 4 (1953).
[2] R. J. Friauf, J. Chem. Phys. **22**, 1329 (1954).
[3] G. Jaffé, Ann. Physik **16**, 217, 249 (1933); Phys. Rev. **85**, 354 (1952).
[4] S. Nudelman and F. Matossi, J. Electrochem. Soc. **101**, 546 (1954).
[5] D. Curie, J. phys. radium **14**, 510, 672 (1953).
[6] S. Nudelman (private communication).
[7] Homer, Rulon, and Butler, J. Electrochem. Soc. **100**, 566 (1953).

tions to traps are not included. These are not important in electroluminescence, as indicated by the much faster decay with field excitation compared to that with ultraviolet excitation of the green luminescence band.[4]

The equations can be solved approximatively for sinusoidal voltages $V=\int E dx=V_1 \exp(i\omega t)$ by writing for n, p, n_c, and E expressions of the form[1]

$$n(x,t)=n_0(x)+n_1(x)\exp(i\omega t) \tag{2}$$

and neglecting terms with higher powers of $\exp(i\omega t)$. Furthermore, it is assumed

$$n_0=p_0=c_0, \tag{3}$$

which implies $E_0=0$, $(n_c)_0=N-c_0$, where $N=$ total number of centers.

These assumptions restrict somewhat the applicability of the equations. This is discussed in detail by MacDonald.[1] The restrictions do not affect, however, the general features of the solutions as long as these are essentially determined by n_1, p_1, $(n_c)_1$, E_1. The changes in time will be described with sufficient approximation even if correct solutions for equilibrium may not be obtained. The assumption of bimolecular recombinations is also an approximation, which is usual in considerations of this type. It is justified in cases of continuous excitation, as is the case in electroluminescence, even if other processes should prevail.[8] Thus, increased qualitative insight is expected in applying the equations to electroluminescence effects even if it may not be possible to use the final formulas for quantitatively exact evaluation.

The general solution of (1) is then given by

$$\begin{aligned} n_1&=A^+\sinh[\rho^+(x-L/2)]+A^-\sinh[\rho^-(x-L/2)],\\ p_1&=B^+\sinh[\rho^+(x-L/2)]+B^-\sinh[\rho^-(x-L/2)], \end{aligned} \tag{4}$$

where L is the thickness of semiconductor layer; ρ^+ and ρ^- are given by the coefficients of the differential equations for p_1 and n_1, and the ratios of the constants A^+, etc., can be determined from the boundary conditions, for which the condition is accepted that charges cannot flow across the surfaces of the phosphor particles.

For the application to electroluminescence, a term which describes the effect of excitation of centers by the periodic field is added to the terms considered in Eq. (1). If k_1 is replaced by

$$k_1=k_0+k'\exp(i\omega t), \tag{5}$$

the equations for p_1 and n_1 can be written, after some elementary calculations, in the form

$$\begin{aligned} i\omega p_1&=(k_0\lambda-k_2c_0)(p_1+n_1)\\ &\quad+D_p\partial^2p_1/\partial x^2-\mu_pc_0\beta(p_1-n_1)\\ &\quad+k'[N-c_0+k_0\lambda(c_0-N)/k_2c_0],\\ i\omega n_1&=(k_0\lambda-k_2c_0)(p_1+n_1)\\ &\quad+D_n\partial^2n_1/\partial x^2+\mu_nc_0\beta(p_1-n_1)\\ &\quad+k'[N-c_0+k_0\lambda(c_0-N)/k_2c_0],\\ \lambda&=k_2c_0/(k_0+i\omega),\quad \beta=4\pi e/\epsilon. \end{aligned} \tag{6}$$

The last term can be simplified by assuming weak excitation so that $N\gg c_0$.

Equation (5) separates a steady excitation from a periodic one. k_0 is interpreted as describing excitation by ultraviolet radiation, and k', as the contribution of the field to excitation. Since the field is described by $E_1\exp(i\omega t)$, all the effects related to the field have the period 2π, too, and also zero average values. However, the observed electroluminescence effects are, of course, always positive and have, furthermore, the period π, since a negative field has the same effect as a positive one. This purely formal difference is a consequence of the aforementioned assumptions and is not essential to the following considerations, since only the changes produced by the field are of interest. The frequency of the effective field, which is twice the frequency of the external field, will therefore be considered as the frequency ω of the equations, notwithstanding the fact that some phosphor particles or parts of them may emit only in the one half-period and not in the other one.[9] This is due to special conditions at the surface of the particles which restrict the actions of the effective field essentially to the one half-period. This need not be considered in the mathematical formalism.

Equation (5) furthermore assumes that the excitation follows the field. This implies that E must be understood as the internal field in the phosphor particle, not as the applied field. Any phase shifts considered here are, therefore, measured relative to this internal field, which itself may be shifted with respect to the applied field. The internal field can be obtained from the last Eq. (1).

The calculations can be substantially simplified by assuming $\mu_p=\mu_n$ and $D_p=D_n$, although $\mu_p\neq\mu_n$ would probably be a more realistic description of the properties of the phosphor particles. The general rules will not be affected by this simplification, so that only the case of equal mobilities will be treated.

For $\mu_p=\mu_n=\mu$, $D_p=D_n=D$, we have

$$\begin{aligned} \partial^2(p_1-n_1)/\partial x^2&=a(p_1-n_1),\\ \partial^2(p_1+n_1)/\partial x^2&=b(p_1+n_1)+c, \end{aligned} \tag{7}$$

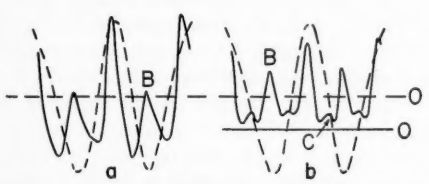


FIG. 1. Brightness waves, 60 cps. *a*: with, *b*: without ultraviolet. ——— Light output, – – – – field. The zero level of light output in *a* is far below the lower edge of the figure (superposition of ultraviolet excited steady output). The units of light output are different for *a* and *b*.

[8] D. Curie, Ann. phys. **7**, 746 (1952).

[9] J. F. Waymouth, Phys. Rev. **95**, 941 (1954).

with

$$aD=2\mu c_0\beta+i\omega, \quad bD=2k_2c_0-2k_0\lambda+i\omega,$$
$$cD=2k'N(\omega^2+ik_0\omega)/(k_0^2+\omega^2).$$

The solutions are

$$p_1=p_{10}-c/2b, \quad n_1=n_{10}-c/2b, \tag{8}$$

where p_{10} and n_{10} are the solutions of Eq. (7) for $k'=0$ or $c=0$, i.e.,

$$p_{10}=-n_{10}=A\sinh\alpha(x-L/2) \tag{9}$$

with $\alpha=2(\mu c_0\beta/2D)^{\frac{1}{2}}(u+iv)$; u^2, $v^2=\frac{1}{2}[(1+\omega^2/4\mu^2c_0^2\beta^2)^{\frac{1}{2}}\pm1]$.

The intensity is given by

$$I=\text{const}\,pn$$
$$=\text{const}[c_0^2+c_0(p_1+n_1)\exp(i\omega t)+p_1n_1\exp(2i\omega t)],$$

which can be written, because of Eqs. (8) and (9), as

$$I=\text{const}[c_0^2-c_0(c/b)\exp(i\omega t)$$
$$-(p_{10}^2-c^2/4b^2)\exp(2i\omega t)]. \tag{10}$$

We write

$$-c/b=\sigma+i\tau, \quad p_{10}^2=\xi+i\eta,$$

where

$$\sigma=(1/M)2k'N(2k_2c_0+k_0), \quad \tau=-(1/M)2k'N\omega,$$
$$M=4k_2^2c_0^2+4k_0k_2c_0+k_0^2+\omega^2. \tag{11}$$

ξ and η can be determined from Eq. (9). In order to eliminate the dependence on x, an average value $\langle p_{10}^2\rangle_{Av}$ may be used. The exact expressions for ξ and η are not important. It is sufficient to know that ξ and η grow with frequency.

Equation (10) can be interpreted as representing a brightness wave with two components U and V of periods 2π and π, respectively, superimposed on a constant background. It is

$$U=c_0(\sigma^2+\tau^2)^{\frac{1}{2}}\exp(i\omega t)=U_0\exp(i\omega t),$$

$$V=\left[\left(\frac{\sigma^2-\tau^2}{4}-\xi\right)^2+\left(\frac{\sigma\tau}{2}-\eta\right)^2\right]^{\frac{1}{2}}\exp(2i\omega t)$$
$$=V_0\exp(2i\omega t). \tag{12}$$

These components have phase shifts against $\exp(i\omega t)$ or $\exp(2i\omega t)$, φ_U and φ_V, given by

$$\tan\varphi_U=-\tau/\sigma=\omega/(2k_2c_0+k_0),$$

$$\tan\varphi_V=\left(\eta-\frac{\sigma\tau}{2}\right)\Big/\left(\frac{\sigma^2-\tau^2}{4}-\xi\right). \tag{13}$$

If ξ and η can be neglected, i.e., for very low frequencies and very weak excitation, we have

$$\varphi_V=2\varphi_U \quad \text{and} \quad V_0=U_0^2/c_0.$$

In this case, the superposition of the two components leads to two maxima coinciding with the maxima and minima of the U-component.

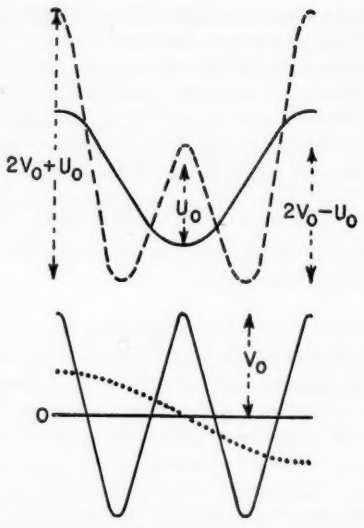


FIG. 2. Components of brightness waves. —— U and V. - - - - Superposition of U and V. ···· Applied field. The minimum of $U+V$ is not exactly at the place of the minimum of V; this is neglected in Eq. (14).

As can be seen from Fig. 2, the heights of these maxima are approximately given by $I_B=2V_0+U_0$, $I_C=2V_0-U_0$, respectively, so that the ratio of the intensities is

$$\rho=I_C/I_B=(2V_0-U_0)/(2V_0+U_0). \tag{14}$$

If ξ and η cannot be neglected, so that $\varphi_V\neq2\varphi_U$, Eq. (14) should be modified, but, even as it stands, it provides a reasonable approximation also in the more general case. In Fig. 2, we have assumed $2V_0>U_0$, which will be true for sufficiently weak excitation.

DISCUSSION

Because of the very complicated dependence of ρ on the parameters ω, c_0, k_0, k', k_2, μ of the problem, only a very general discussion can be given; but this is sufficient for the purpose of a qualitative comparison with the observations.

I_B and I_C are identified with the observed B- and C-peaks. It is easily seen that $U_0/2V_0$ becomes smaller with increasing frequency as soon as the contribution of the ξ and η terms becomes appreciable; that is, for not too low frequencies (if ω is not small compared to $\mu c_0\beta$). Therefore, the observed increase of the C-peaks with respect to the B-peaks can be expected since ρ tends to unity if $U_0/2V_0$ becomes zero. Low mobilities and weak excitation are essential for the observation of this phenomenon. Although numerical values are not available, this conclusion seems to be compatible with the general experience.

If k_0 increases, $U_0/2V_0$ increases too. This is obvious for the case of small values of ξ and η, but it will be true also for higher values because the dependence on

k_0 is contained only in σ and τ. Therefore, ρ becomes smaller, indicating that steady excitation by ultraviolet suppresses the C-peak.

Since this result has been obtained in a rather involved manner, a nonmathematical consideration may make it more plausible. Without ultraviolet, two processes superpose. One is the excitation of electrons by the field and their subsequent recombination (B-peaks). The other is movement of electrons under the influence of the field producing inhomogeneous distribution of charges. When and if these charges diffuse back, they too may recombine (C-peaks). Ultraviolet radiation increases the number of electrons excited to the conduction band, which tends to decrease the charge inhomogeneity and, therefore, decreases the chance of recombination through back diffusion. The fact that the C-peaks recover only slowly after cutting off the ultraviolet excitation may, then, be interpreted as indicating that the electrons produced by the ultraviolet radiation disappear only gradually so that the case of a pure polarization effect is attained only after some time. That these electrons disappear more slowly than would be expected from the decay of ultraviolet excited phosphorescence, may be understood by the consideration that the field has displaced the electrons from the centers so that recombination is more difficult.

The case of absence of polarization charges may be studied by considering Eq. (1) again, but for $\mu=0$, $D=0$. This condition prevents the piling up of charges. Then the equations reduce to

$$\partial p/\partial t=\partial n/\partial t=[k_0+k' \exp(i\omega t)]N-k_2p^2, \quad (15)$$

if again weak excitation is assumed. This is an equation of Mathieu's type, which admits periodic solutions only for particular values of the parameters.[10] Although this equation may, therefore, not be useful for the stationary state, it can be used for the nonstationary state in the first instants after the application of the field, when polarization has not yet come into play. It is convenient to modify Eq. (15) slightly by writing, for $k_0=0$,

$$\partial p/\partial t=k' \cos^2\omega t-k_2p^2.$$

$\cos^2\omega t$ is an approximation for $|\cos\omega t|$. This takes into account the equivalence of positive and negative fields. A possible solution is

$$I=(\omega^2/k_2)(\zeta+\phi'/\phi)^2 \tanh^2\zeta\omega t, \quad (16)$$

where ϕ is a periodic function with the period π and the properties $\phi(-x)=\phi(x)$, $\phi(-x)'=-\phi(x)'$ or $\phi(-x)=-\phi(x)$, $\phi(-x)'=\phi(x)'$; $\phi(x)'=d\phi(x)/dx$, $x=\omega t$. ζ is a parameter, approximatively given by $\zeta^2=k_2k'/2\omega^2$. This represents a periodic ripple with increasing amplitude superimposed on a steady rise of average intensity. This corresponds well to observations.[11]

For long times, Eq. (16) describes, of course, a periodic behavior of the brightness waves, but specific statements about phases or number of maxima cannot be made. In any case, the general theory of Mathieu's equation permits only one linearly independent periodic solution if there is one at all.

Finally, it may be pointed out that the phase shifts indicated by Eq. (13) cannot be observed directly since additional frequency-dependent phase shifts may occur that are also due to polarization effects, viz., to the phase shift of the internal field with respect to the applied field. For the problem of this paper, these phase shifts are irrelevant. It shall only be mentioned that the expression for the phase shift of the fields derived from Eqs. (1) and (9) is different from that given by Destriau and Curie[5] who take into account only the continuity of charge flow and the Poisson equation.

There are still other phase shifts observed[6] if small fields are applied to ultraviolet excited electroluminescent phosphors. An extinguishing effect is obtained, and the brightness waves are shifted in the opposite direction than in the case of high fields. This effect is due to changes in the number of available electrons, not to moving polarization charges, and is controlled by the lifetime of the electrons in the conduction band.[5]

CONCLUSION

The foregoing considerations sufficiently indicate that the processes occurring in semiconductors as inferred from capacitance effects are also responsible for electroluminescence effects. Moving polarization charges are essential for the description of the observed characteristic details of brightness waves.

[10] E. T. Whittaker and G. N. Watson, *Modern Analysis* (University Press, Cambridge, 1927); L. Brillouin, *Wave Propagation in Periodic Structures* (McGraw-Hill Book Company, Inc., New York, 1946).

[11] F. Vigean, Compt. rend. **236**, 1151 (1953).

Magnetic Rotation Phenomena in a Polycrystalline Ferrite. II*

DAVID PARK
Williams College, Williamstown, Massachusetts
(Received December 27, 1954)

The calculations of an earlier paper are extended to the determination of the entire permeability tensor of a polycrystalline sample of a ferrimagnetic material in which there is a slight over-all magnetization. The calculation takes account of the largely random orientation of the magnetic axes of the crystallites and of the fact that their characteristic resonant frequencies are statistically distributed. It is shown that the magnetic interaction between crystallites can be included, but it does not greatly affect the form of the final results.

1. INTRODUCTION

IN an earlier communication[1] we have studied the high-frequency magnetic properties of a sintered polycrystalline sample of ferrimagnetic material, assuming that the dispersion is caused by magnetic rotation and that the crystallites differ in their characteristic resonant frequencies and in the orientation of their magnetic axes, which was assumed random. In the work to be reported here we have considered the situation in which the sample shows a net uniform magnetization. The magnetized crystallites are no longer oriented at random, but in some axially symmetric distribution to be specified later. The remaining assumptions will be the same as in I. The permeability is now a tensor, and whereas in I the off-diagonal terms averaged out, here they will give rise to a magneto-optical rotation (the "ferromagnetic Faraday effect[2]"). It is hoped that these results can help in finding out about the mechanisms at work in ferrimagnetic substances. Progress in this direction has been hampered by the fact that calculations of the diagonal part of the permeability tensor based on two very different assumptions—that the high-frequency permeability is due to Bloch wall motion and that it is due to domain rotation—have both given results at least roughly in agreement with experiment.[3] It appears unlikely, however, that wall motion can produce a Faraday effect comparable to that produced by domain rotation. We shall show below that the rotatory properties of a sample of any shape are completely determined in the theory of domain rotation when one knows the spectrum of the complex initial permeability. We have treated the initial permeability (i.e., the situation in which there is essentially no externally applied magnetizing field) because it seems that here the distinction between the two theories will be the most critical.

The calculation will be given in three steps, which consist of finding the permeability tensor μ successively in an infinite homogeneous medium, an infinite polycrystalline medium, and a polycrystalline sample of finite extent. All the main results of I will be rederived here as special cases.

2. PERMEABILITY OF AN INFINITE CRYSTAL

We start with the equation of motion of $\mathbf{M}$, the magnetic moment per unit volume, in the form first given by Landau and Lifschitz[4]:

$$\partial\mathbf{M}/\partial t=\gamma\mathbf{M}\times\mathbf{H}_T-(\lambda/M^2)\mathbf{M}\times\mathbf{M}\times\mathbf{H}_T, \quad (1)$$

where $\mathbf{H}_T$, composed of the alternating applied field $\mathbf{h}$ and the fixed anisotropy field $\mathbf{H}_{A0}$, is the total field acting on $\mathbf{M}$, γ is the effective value of the gyromagnetic ratio of the two-lattice system,[5] and λ is a constant responsible for the damping. Setting M_x+iM_y equal to M_+ and similarly for the other quantities, we find that (1) can be written

$$\dot{M}_+=-i\gamma'(H_{A0}M_+-h_+M_s), \quad (2)$$

where

$$\gamma'=(1-i\alpha)\gamma, \quad \alpha=\lambda/\gamma M_s \quad (3)$$

and we have chosen the axes so that the z-axis is parallel with $\mathbf{H}_{A0}$. We have written M_s for the approximately constant z-component of the magnetization, and have assumed that the z-component of $\mathbf{h}$ is negligible beside H_{A0}. If we assume that all fluctuating quantities are proportional to $e^{-i\omega t}$ (noting as in I that the two imaginaries i and j are to be distinguished), we can solve (2) to find for the alternating susceptibility:

$$\chi_+=M_+/h_+=\frac{\gamma M_s}{(ij\omega/1-i\alpha)+\gamma H_{A0}}. \quad (4)$$

It is customary to express the relation between the alternating quantities $\mathbf{b}$ and $\mathbf{h}$ in terms of a permeability tensor[6]

$$\mu=\begin{pmatrix}\mu & j\kappa\\ -j\kappa & \mu\end{pmatrix} \quad (5)$$

by writing

$$\begin{pmatrix}b_x\\ b_y\end{pmatrix}=\mu\begin{pmatrix}h_x\\ h_y\end{pmatrix}. \quad (6)$$

* This work was supported by the Sprague Electric Company.
[1] D. Park, Phys. Rev. **97**, 60 (1955), cited as I.
[2] C. L. Hogan, Bell System Tech. J. **31**, 1 (1952); Revs. Modern Phys. **25**, 253 (1953).
[3] G. T. Rado, Revs. Modern Phys. **25**, 81 (1953).
[4] L. Landau and E. Lifschitz, Physik. Z. Sowetjunion **8**, 153 (1935).
[5] R. K. Wangsness, Phys. Rev. **93**, 68 (1954); F. Brown and D. Park, Phys. Rev. **93**, 381 (1954).
[6] D. Polder, Phil. Mag. **40**, 99 (1949). We use the opposite sign of j from Polder and Hogan.

The same relation can also conveniently be written

$$b_+=(\mu-ij\kappa)h_+=(1+4\pi\chi_+)h_+, \tag{7}$$

whence we have (in the notation of **I**):

$$\mu-1=4\pi\,\mathrm{Re}_i\chi_+, \quad \kappa=4\pi j\,\mathrm{Im}_i\chi_+. \tag{8}$$

The advantage of using the variables labeled + and − is that we have thereby diagonalized the permeability tensor.

3. PERMEABILITY OF AN INFINITE POLYCRYSTALLINE MEDIUM

In order to analyze the magnetization of a uniformly magnetized polycrystalline medium, let us focus attention on a crystallite X, surrounded by a group of neighbors. We shall assume that the neighbors are typical enough so that their average alternating magnetization is close to **M**, the average of the sample as a whole. The total alternating field acting on the crystallite X is, as in **I**, given roughly by an expression of the form

$$\mathbf{H}=\mathbf{h}-\rho\mathbf{M}, \tag{9}$$

where the second term is to take account of the interaction between crystallites. We shall assume that any external magnetizing field present is much weaker than H_{A0}, so that it does not affect the resonant frequencies of the crystallites. To take it into account is straightforward, unless the external field happens to be comparable to H_{A0}, but it complicates the analysis.

We shall need to use two systems of coordinates, a fixed system $\mathfrak{K}$ in which the z-axis is the direction of the static magnetization of the whole substance, and a system $\mathfrak{X}$ in which the z-axis is parallel to the static magnetization of X. In this latter system we have

$$M_+=\chi_+H_+, \quad M_-=\chi_-H_-, \tag{10}$$

where, as in (7),

$$\chi_+=(1/4\pi)(\mu-1-ij\kappa), \quad \chi_-=(1/4\pi)(\mu-1+ij\kappa). \tag{11}$$

Suppose that $\mathfrak{A}$ is a matrix which rotates $\mathfrak{X}$ into $\mathfrak{K}$. Then, if **H** is given in $\mathfrak{K}$, it will be $\mathfrak{A}^{-1}\mathbf{H}$ in $\mathfrak{X}$. The alternating magnetization will be given in $\mathfrak{X}$ by $\chi\mathfrak{A}^{-1}\mathbf{H}$, and in $\mathfrak{K}$ this becomes

$$\mathbf{M}=\mathfrak{A}\chi\mathfrak{A}^{-1}\mathbf{H}. \tag{12}$$

We can omit the z-components of both **H** and **M**, the former by our choice of the system $\mathfrak{K}$ and the latter because it will clearly vanish by symmetry. Then if $\alpha, \beta, \cdots$ stand for + or −, (12) is

$$M_\alpha=\mathfrak{A}_{\alpha\beta}\chi_\beta\mathfrak{A}^{-1}_{\beta\gamma}H_\gamma. \tag{13}$$

In terms of a standard set of Euler angles,[7] we have[8]

$$\mathfrak{A}=\begin{pmatrix} a^2 & -b^2 \\ -b^{*2} & a^{*2} \end{pmatrix}, \quad \mathfrak{A}^{-1}=\begin{pmatrix} a^{*2} & -b^2 \\ -b^{*2} & a^2 \end{pmatrix},$$

where

$$a=e^{-\frac{1}{2}i(\psi+\varphi)}\cos\tfrac{1}{2}\theta, \quad b=-ie^{-\frac{1}{2}i(\psi-\varphi)}\sin\tfrac{1}{2}\theta,$$

θ being taken as the angle between the two z-directions. With this we can write out (13). Keeping only the terms which do not vanish when averaged over all φ orientations of X, we find for the alternating components

$$\begin{aligned} 4\pi M_+&=[(\mu-1-ij\kappa)\cos^4\tfrac{1}{2}\theta \\ &\qquad +(\mu-1+ij\kappa)\sin^4\tfrac{1}{2}\theta]H_+, \\ 4\pi M_-&=[(\mu-1-ij\kappa)\sin^4\tfrac{1}{2}\theta \\ &\qquad +(\mu-1+ij\kappa)\cos^4\tfrac{1}{2}\theta]H_-, \end{aligned} \tag{14}$$

and for the fixed component we have evidently

$$M_z=M_s\cos\theta, \tag{15}$$

where M_s is the saturation magnetization of the crystallite. Averaged over all θ, these become

$$4\pi\langle M_\pm\rangle=\tfrac{2}{3}(\mu-1)H_\pm, \quad \langle M_z\rangle=0, \tag{16}$$

as was found by another way in **I**.

If there is a net magnetization the average over θ becomes a weighted average. Let us suppose that far from saturation the weighting function is

$$f(\theta)=\tfrac{1}{2}(1+k\cos\theta) \quad (k<1).$$

With this we find that the average $\langle M_z\rangle_f$ becomes $\frac{1}{3}kM_s$. This is the net magnetization of the sample, which we shall call $\bar{M}$ and in terms of it we can write the remaining averages as

$$4\pi\langle M_\pm\rangle_f=[\tfrac{2}{3}(\mu-1)\mp ij\kappa\bar{M}/M_s]H_\pm, \tag{17}$$

with μ and κ given by (8). Another choice of angular distribution function will introduce slightly different numerical coefficients, but the difference should be unimportant.

We can now carry out the average over the different characteristic resonant frequencies. Assuming as in **I** a frequency distribution function

$$q(\omega_x)=\frac{1}{\pi(A-B)}\left(\tan^{-1}\frac{A-\omega_x}{\sigma_1}-\tan^{-1}\frac{B-\omega_x}{\sigma_2}\right), \tag{18}$$

where we have denoted γH_{A0} in (4) by ω_x, we have for the averaged values

$$\langle\mu\rangle_q-1=4\pi\gamma M_s\,\mathrm{Re}_i p(\omega), \quad \langle\kappa\rangle_q=4\pi j\gamma M_s\,\mathrm{Im}_i p(\omega), \tag{19}$$

where

$$p(\omega)\cong\int_{-\infty}^{\infty}\frac{q(\omega_x)d\omega_x}{(ij\omega/1-i\alpha)+\omega_x}.$$

The integrals can be evaluated exactly, and writing

[7] H. Goldstein, *Classical Mechanics* (Addison-Wesley Press, Cambridge, 1951).

[8] See the Appendix.

$\langle\mu\rangle_q=\langle\mu'\rangle_q+j\langle\mu''\rangle_q$, etc., we find

$$\langle\mu'\rangle_q=1+\pi\frac{\gamma M_s}{A-B}\ln\frac{(\omega^2-A^2+\sigma_1^2)^2+4A^2\sigma_1^2}{(\omega^2-B^2+\sigma_2^2)^2+4B^2\sigma_2^2}, \tag{20a}$$

$$\langle\mu''\rangle_q=2\pi\frac{\gamma M_s}{A-B}\left(\tan^{-1}\frac{2A\sigma_1}{\omega^2-A^2+\sigma_1^2}-\tan^{-1}\frac{2B\sigma_2}{\omega^2-B^2+\sigma_2^2}\right), \tag{20b}$$

$$\langle\kappa'\rangle_q=\pi\frac{\gamma M_s}{A-B}\ln\frac{[(A-\omega)^2+\sigma_1^2][(B+\omega)^2+\sigma_2^2]}{[(A+\omega)^2+\sigma_1^2][(B-\omega)^2+\sigma_2^2]}, \tag{20c}$$

$$\langle\kappa''\rangle_q=2\pi\frac{\gamma M_s}{A-B}\left(\tan^{-1}\frac{2\sigma_1\omega}{\omega^2-A^2-\sigma_1^2}-\tan^{-1}\frac{2\sigma_2\omega}{\omega^2-B^2-\sigma_2^2}\right). \tag{20d}$$

Finally, we can write down the formulas which take into account the interaction between crystallites although, for reasons set forth in **I**, we have felt it unnecessary to use them in computation. If we average (17) over ω_x and write the result as $\langle M_\pm\rangle_{f,q}=\langle\chi_\pm\rangle_{f,q}H_\pm$ and define the actual susceptibility of the substance as $\chi_{a+}=\langle M_+\rangle_{f,q}/h_+$, we can use (9) to find

$$\chi_{a+}=\frac{\langle\chi_+\rangle_{f,q}}{1+\rho\langle\chi_+\rangle_{f,q}},$$

of which (**I** 13) is a special case. Neglecting ρ, we have for the μ and κ of the bulk substance

$$\mu_a-1=\tfrac{2}{3}(\langle\mu\rangle_q-1),\quad \kappa_a=\bar{M}\langle\kappa\rangle_q/M_s. \tag{21}$$

Figures 1 and 2 show the spectra of μ_a and κ_a which are computed from the constants deduced for a particular sample of nickel ferrite in **I**. The frequency at which κ_a' vanishes is given in general by

$$\omega_0\approx[B(A^2+\sigma_1^2)/A]^{\frac{1}{2}}. \tag{22}$$

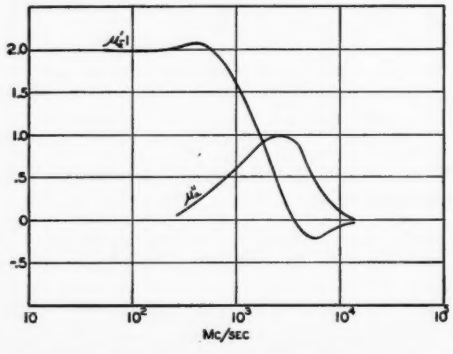


FIG. 1. Real and imaginary parts of the diagonal element μ_a of the bulk permeability tensor computed for a particular sample of nickel ferrite.

The predicted values of κ_a are smaller than those plotted by a factor of $3\bar{M}/2M_s$, assumed to be much less than unity. (The factor of $\frac{2}{3}$ was introduced into κ_a so that corresponding functions in (20) would appear on the same scale.) It will be noted that κ_a is linearly proportional to ω at low frequencies. The relation as given by (20) and (21) is

$$\kappa_a\approx\frac{4\pi\gamma\bar{M}\omega}{A-B}\left(\frac{1}{B+j\sigma_2}-\frac{1}{A+j\sigma_1}\right). \tag{23a}$$

In the case discussed in **I**, in which $\sigma_1/A=\sigma_2/B=\lambda$, we find for the first two terms

$$\kappa_a\approx\frac{4\pi\gamma\bar{M}\omega}{(1+j\lambda)AB}\left(1+\tfrac{1}{3}\left(\frac{\omega}{1+j\lambda}\right)^2\frac{A^2+AB+B^2}{A^2B^2}\right). \tag{23b}$$

These constants describe the behavior of the material in bulk. For experimental purposes, however, it is often

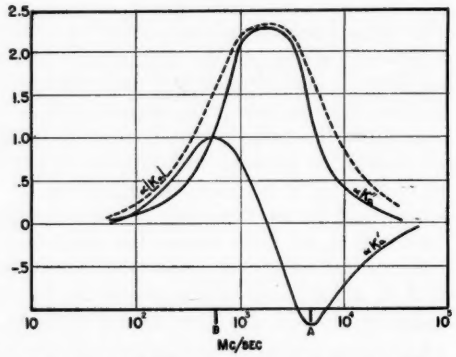


FIG. 2. Real and imaginary parts and modulus of the off-diagonal element κ_a. The ordinate is $\alpha\kappa_a$, where α is $2M_s/3\bar{M}$.

convenient to have the ferrite in the form of a small sample, sometimes several small spheres, next to the short-circuited end of a wave guide, and it has been pointed out by Brown[9] that spectra measured in this way may differ markedly from those predicted by (20) because of the influence of demagnetizing factors. What follows is essentially Brown's analysis, extended to take account of the off-diagonal components of $\boldsymbol{\mu}$.

4. EFFECTIVE PERMEABILITY

Let $\mathbf{b}$ be the alternating induction within the sample, and define an effective permeability tensor $\boldsymbol{\mu}_e$ by

$$\mathbf{b}=\boldsymbol{\mu}_e\mathbf{h}, \tag{24}$$

where $\mathbf{h}$ is the applied external field. The field within the sample is

$$\mathbf{h}_i=\mathbf{h}-\mathbf{NM},$$

[9] F. Brown (unpublished). See also C. Kittel, Phys. Rev. **73**, 155 (1948).

where $\mathbf{N}$ is the demagnetizing tensor and $\mathbf{M}$ is the magnetization. We have further:

$$\mathbf{b}=\mathbf{h}_i+4\pi\mathbf{M} \quad \text{and} \quad \mathbf{b}=\mathfrak{y}_a\mathbf{h}_i,$$

where $\mathfrak{y}_a$ is the actual permeability of the bulk substance as given by (21). From the last four equations it is easy to show that $\mathfrak{y}_e$ and $\mathfrak{y}_a$ are related by

$$\mathfrak{y}_e=\mathfrak{y}_a[1+\mathbf{N}(\mathfrak{y}_a-1)/4\pi]^{-1}. \tag{25}$$

Suppose now that $\mathbf{N}$ is diagonal, and equal to the unit matrix times a number $4\pi\nu$. (The general case is also possible, but the result is rather academic.) Then expressing $\mathfrak{y}_a$ in the form (5), with a similar expression for $\mathfrak{y}_e$, we find after an elementary calculation:

$$\mu_e=\frac{(\mu_a^2-\kappa_a^2)[(1-\nu)\mu_a+\nu(\mu_a^2-\kappa_a^2)]}{[(1-\nu)\mu_a+\nu(\mu_a^2-\kappa_a^2)]^2-(1-\nu)^2\kappa_a^2}, \tag{26a}$$

$$\kappa_e=\frac{(1-\nu)(\mu_a^2-\kappa_a^2)\kappa_a}{[(1-\nu)\mu_a+\nu(\mu_a^2-\kappa_a^2)]^2-(1-\nu)^2\kappa_a^2}. \tag{26b}$$

In the initial permeability region or the remanent state, κ_a^2/μ_a^2 is generally very small. If it is neglected, we find

$$\mu_e\approx\frac{\mu_a}{1-\nu+\nu\mu_a}, \quad \kappa_e=\frac{(1-\nu)\kappa_a}{(1-\nu+\nu\mu_a)^2}. \tag{27}$$

Occasionally in practice, it is only convenient to measure μ_e', μ_e'', and $|\kappa_e|$. By (27), these are

$$\begin{aligned}&\mu_e'=[(1-\nu)\mu_a'+\nu(\mu_a'^2+\mu_a''^2)]/D,\\ &\mu_e''=(1-\nu)\mu_a''/D, \quad |\kappa_e|=(1-\nu)|\kappa_a|/D,\\ &D=(1-\nu+\nu\mu_a')^2+\nu^2\mu_a''^2.\end{aligned} \tag{28}$$

Figure 3 shows how demagnetization affects the permeabilities plotted in Figs. 1 and 2.

For frequencies from zero up to the neighborhood of B, we find from (23b):

$$|\kappa_a|=\frac{4\pi\gamma\bar{M}}{AB}\frac{\omega}{(1+\lambda^2)^{\frac{1}{2}}}\times\left[1+\tfrac{1}{2}\frac{1-\lambda^2}{(1+\lambda^2)^2}\omega^2\frac{A^2+AB+B^2}{A^2B^2}+O(\omega^4)\right]. \tag{29}$$

If data are available on $|\kappa_a|$, μ_e', and μ_e'', then it is probably best to rewrite (27) in terms of μ_e' and μ_e'' to get (for spheres)

$$|\kappa_a|=\frac{6|\kappa_e|}{(3-\mu_e')^2+\mu_e''^2}, \tag{30}$$

then calculate $|\kappa_a|$ from the data, and compare the result with (29) or (21).

Measurements of the spectra of some of these quantities have been undertaken by Brown. A determination of $|\kappa_a|$ is reported in the article following.

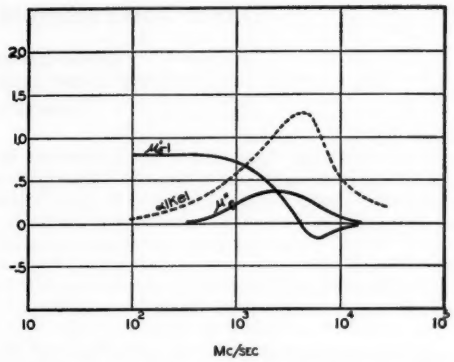


FIG. 3. Elements of the effective permeability tensor for a small spherical sample. Figures 1, 2, and 3 are all drawn to the same scale.

It is a pleasure to acknowledge my indebtedness to Dr. Fielding Brown, who suggested this calculation and has helped greatly in the clarification of the physical problems involved.

APPENDIX

Suppose we have two Cartesian coordinate systems, $\mathfrak{X}'$ and $\mathfrak{R}'$, with variables x, y, z, and a matrix A which rotates $\mathfrak{X}'$ into $\mathfrak{R}'$. If the variables are chosen as x_+, x_-, z, with $x_\pm=x\pm iy$, we shall call the systems $\mathfrak{X}$ and $\mathfrak{R}$. To pass from $\mathfrak{X}'$ to $\mathfrak{X}$ or $\mathfrak{R}'$ to $\mathfrak{R}$ amounts to a transformation of $\mathfrak{X}'$ or $\mathfrak{R}'$ by a matrix

$$C=\begin{pmatrix}1 & i & 0\\ 1 & -i & 0\\ 0 & 0 & 1\end{pmatrix},$$

so that, with respect to $\mathfrak{X}$, the rotation matrix is

$$\mathfrak{A}=CAC^{-1}$$

which with the Euler angles of reference 5 can be written

$$\mathfrak{A}=\begin{pmatrix}a^2 & -b^2 & -2ab\\ -b^{*2} & a^{*2} & -2a^*b^*\\ ab^* & a^*b & aa^*-bb^*\end{pmatrix},$$

with

$$a=e^{-\frac{1}{2}i(\psi+\varphi)}\cos\tfrac{1}{2}\theta, \quad b=-ie^{-\frac{1}{2}i(\psi-\varphi)}\sin\tfrac{1}{2}\theta.$$

To find $\mathfrak{A}^{-1}$, we note that since A is orthogonal, one can write

$$\mathfrak{A}^{-1}=CA^{-1}C^{-1}=C\bar{A}C^{-1}=C\bar{C}\bar{\mathfrak{A}}(C\bar{C})^{-1},$$

where the dash denotes the transpose, and with

$$C\bar{C}=\begin{pmatrix}0 & 2 & 0\\ 2 & 0 & 0\\ 0 & 0 & 1\end{pmatrix}, \quad (C\bar{C})^{-1}=\begin{pmatrix}0 & \frac{1}{2} & 0\\ \frac{1}{2} & 0 & 0\\ 0 & 0 & 1\end{pmatrix},$$

we find

$$\mathfrak{A}^{-1}=\begin{pmatrix}a^{*2} & -b^2 & 2a^*b\\ -b^{*2} & a^2 & 2ab^*\\ -a^*b^* & -ab & aa^*-bb^*\end{pmatrix}.$$

In the restricted calculations reported in this paper, only the submatrix transforming (x_+,x_-) is required.

Direct Observation of Domain Rotation in Ferrites

FIELDING BROWN AND CHARLES L. GRAVEL
Sprague Electric Company, North Adams, Massachusetts
(Received December 27, 1954)

A direct experimental observation of domain rotation in the initial permeability region of sintered nickel ferrite has been made using a technique similar to that originated by F. Bloch to observe nuclear precession. A 5-mm diameter sphere of ferrite is surrounded by crossed transmitter and receiver loops adjusted for minimum rf signal when the sample is demagnetized. Application of an external steady field (as low as 5 oersteds, perpendicular to both loops) then causes a measurable receiver loop output which cannot be eliminated by readjusting the planes of the loops. The motion of the net magnetic moment is thus rotational. In the present case, the external field is needed only to polarize the sphere, the field for Larmor precession being supplied by the equivalent anisotropy fields seen by each individual domain. Since the applied field is much smaller than these anisotropy fields, a plot of the detector loop output (per unit applied field) *vs* frequency delineates the rotational resonance spectrum of the polycrystalline sample. Measurements have been made from 20 Mc/sec to 800 Mc/sec and can be described qualitatively using the theory of domain rotation. In order to understand the relatively low dispersion frequency indicated by the measurements, it is hypothesized that intermediate conditions exist in some polycrystalline ferrites between the situation in which the initial permeability is purely a result of domain rotation and that in which motion of domain walls is the dominant mechanism. The distinction between the models is thus probably not very sharp.

I. INTRODUCTION

THE permeability spectrum of sintered ferrites has been the subject of a number of investigations.[1–6] In all studies previous to that of Rado,[5] the single dispersion of the permeability which was observed in the region of several megacycles to several kilomegacycles was attributed to a natural ferromagnetic resonance in which the domains were considered as precessing in an internal anisotropy field under the influence of external excitation. Wijn *et al.*[7] have also studied the magnetic spectra of ferrites under mechanical stress and have obtained results pointing toward domain rotation as the principal mechanism.

In contrast to this work, Rado has found that the permeability spectrum of Ferramic *A* has two dispersion regions at room temperature, one at about 50 Mc/sec, and the other at 1400 Mc/sec.[8] In this study, the loss peak at lower frequency was ascribed to a resonance of domain walls as suggested by Döring[9] and Becker,[10] whereas the higher frequency one was attributed to domain rotation. This conclusion rested in part on a comparison of the magnetic spectrum of the solid ferrite with the spectrum of the same material measured in powdered form. In the latter case the 50-Mc/sec peak was absent, whereas the 1400-Mc/sec peak remained but in altered form. Rado's identification of the peak at 50 Mc/sec was thus based on the assumption that this resonance, if due to domain rotation in an anisotropy field, would be a fundamental aspect of the material and thus would not be affected by grinding. Its disappearance was thus taken to exclude domain rotation, leaving domain wall motion as the responsible mechanism. Inasmuch as ferromagnetic anisotropy arises from structurally dependent properties (strain and internal demagnetizing fields), as well as from the fundamental magnetocrystalline anisotropy, the argument does not appear entirely convincing. This is especially true in the light of recent work[6,11] in which the magnetic spectrum is found to depend markedly on demagnetizing effects which could certainly be applied either to the sample as a whole or to individual particles within a heterogeneous sample, and which would thus operate differently in a sintered mass and in a powder-wax mixture.

The present writers have studied the magnetic spectrum of nickel ferrite,[6] in an attempt to identify the mechanisms contributing to the initial permeability. As the finely divided and pressed ferrite powder was sintered at increasing maximum temperatures, the single observed loss peak was found to move to lower frequencies. Quantitative comparison of the characteristic frequency of this dispersion with the susceptibility at low frequencies in each case was made using a relation proposed by Snoek[1] and recently analyzed by Park.[12] Agreement between theory and experiment was taken as indicating the existence of domain rotation in the sample studied. Since the polycrystalline grain size was relatively small (0.5 to 3 microns), domain walls were not considered to contribute significantly in the particular comparison made.

The present work originated in an attempt to obtain more direct evidence of the mechanism involved and

[1] J. L. Snoek, Nature **160**, 90 (1947); Physica **14**, 207 (1948).
[2] Brockman, Dowling, and Steneck, Phys. Rev. **77**, 85 (1950).
[3] J. B. Birks, Proc. Phys. Soc. (London) **B63**, 65 (1950).
[4] Welch, Nicks, Fairweather, and Roberts, Phys. Rev. **77**, 403 (1950).
[5] G. T. Rado, Revs. Modern Phys. **25**, 81 (1953); this reference refers to earlier work of Rado and others.
[6] F. Brown and C. L. Gravel, Phys. Rev. **95**, 652 (1954).
[7] Wijn, Gevers, and van der Burgt, Revs. Modern Phys. **25**, 91 (1953).
[8] Principally magnesium ferrite.
[9] W. Döring, Z. Naturforsch. **3a**, 374 (1948).
[10] R. Becker, Physik. Z. **39**, 856 (1938); Ann. Physik (5) **36**, 340 (1939).
[11] D. Park, preceding paper [Phys. Rev. **98**, 438 (1955)] in which a derivation is given of effective values of κ_e and μ_e in terms of their values for bulk material. This paper is hereafter referred to as II.
[12] D. Park, Phys. Rev. **97**, 438 (1955).
[13] Bloch, Hansen, and Packard, Phys. Rev. **70**, 474 (1946).

makes use of a technique similar to the first used by Bloch[13] to observe nuclear precession. By arranging two mutually perpendicular loops around a small sphere of ferrite, an attempt has been made to observe the induction signal which results from precession of the net ferrimagnetic moment in an external steady field. Inasmuch as it is possible to show that wall motions, at least as normally visualized in a multidomain sample could produce no appreciable signal under these circumstances, such a device may be used to demonstrate and measure domain rotation in ferrites.

In the present case the induction signal results from a superposition of individual signals from domains, each of which can be thought to precess independently within its own effective anisotropy field. When a small external steady field is applied to the sample, the magnetic moments of the domains rotate slightly in the field direction giving rise to a magnetic moment for the sample as a whole. However, for small steady fields, the effective field experienced by individual domains does not depart appreciably from its zero external field value, and the natural ferrimagnetic resonance frequencies of the domains remain essentially unaltered. The external field is here employed only to give sufficient polarization to the sample to permit measurement. Under these conditions the signal induced in the secondary loop can be expected to be proportional to the net moment, and thus to the size of the external steady field (for small fields), and is also a measure of the extent to which the net magnetic moment departs from the steady field direction during precession. If the initial susceptibility of the ferrite throughout the radio-frequency spectrum is principally due to domain rotations, the secondary output per unit steady field should be connected in some understandable way with the observed magnetic permeability spectrum of the material. The present work is an attempt to observe this connection.

As to whether or not domain wall motion may give a measurable output from such a device and thus yield ambiguity in interpretation, it is important to notice that for a polycrystalline sample such as that here employed, the individual domains are oriented in random directions, and it can be expected that, if true domain walls actually exist, clockwise and counter-clockwise spin distributions will be found in equal numbers in traversing the walls between domains. Thus even though wall motion may be gyroscopic in nature, the bulk material is isotropic, since for every domain wall oriented in such a way that the composite spins may contribute to the induction signal, there will exist another wall which will induce an equal and opposite signal. It is also expected that the material will respond only in the signal direction, i.e., be nonrotational, even under conditions of partial magnetic saturation, since the only effect of a small steady field would be to move existing domain walls without a net alteration in their number or change in the balance between

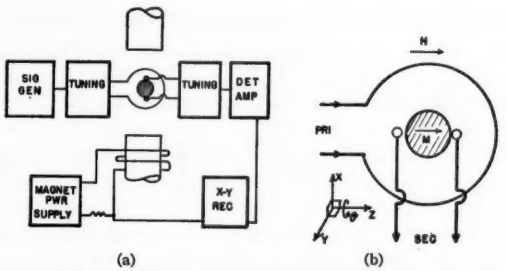


FIG. 1. (a) Block diagram of apparatus for observing domain rotation. (b) Diagrammatic sketch of rf head.

clockwise and counter-clockwise spin distributions. Therefore even under these conditions the precession of spins within moving walls cannot contribute to the output. The possibility of a signal contribution arising from special conditions of crystalline orientation, strain, inclusions, shapes, or other inhomogeneities which might be expected to affect wall motion can be ruled out, since the present sample may contain 10^9 to 10^{10} domains, assuming a single domain to encompass only one or, at most, two to three crystallites. Microscopic anisotropies are therefore also expected to statistically cancel for the bulk sample.

II. EXPERIMENTAL

A. Apparatus

The crossed-coil ferrimagnetic induction device is represented schematically in Fig. 1(b). The primary and secondary coils are normally in the x-z and y-z directions, respectively, and the steady field H_0, used to produce the net magnetic moment, M, is in the z-direction. If it is assumed that domain rotation occurs within the spherical sample, the magnetic moments of domains will each precess in response to an applied alternating field h_y, with the result that M will also precess and induce a secondary signal. To connect the signal with the permeability spectrum of the ferrite thus requires measurement over the entire frequency range of the dispersion.

In the case of nickel ferrite the dispersion occurs between approximately 20 Mc/sec and several thousand Mc/sec, so that it is necessary to make measurements at frequencies where voltages are difficult to determine. This situation is further aggravated by the fact that *effective* values of the permeability components are actually measured, as opposed to values characteristic of infinite media, the demagnetizing effect here involved causing the dispersion region of the material to be removed to much higher frequencies.[6,11] We have therefore used the microwave practice of arranging the measurement in such a way that results are obtained as voltage ratios. This can readily be seen from the following theory of measurement.

As has been shown,[14] a ferrimagnetic material under

[14] D. Polder, Phil. Mag. 40, 99–115 (1949).

conditions of magnetic polarization obeys the relation

$$b_i=\mu_{ik}h_k, \tag{1}$$

where the tensor, μ_{ik}, has the components

$$\mu_{ik}=\begin{bmatrix} \mu & j\kappa \\ -j\kappa & \mu \end{bmatrix}, \tag{2}$$

when the net magnetic moment of the bulk sample, M, is in the z-direction and h is applied in the x-y plane.

In Fig. 1(b), if transmitter and receiver loops are at right angles to each other and parallel to the coordinate planes, Eq. (1) may be written

$$b_x=j\kappa_e h_{y0}, \tag{3a}$$

$$b_y=\mu_e h_{y0}, \tag{3b}$$

where μ_e and κ_e are effective values of the permeability components appropriate to the spherical sample, and h_{y0} is the applied alternating field exterior to the sphere. Under these conditions, the voltage induced in the receiver loop is obtained from (3a) as

$$E_1=SMnA\omega h_{y0}\kappa_{e1}, \tag{4}$$

where n is the number of secondary turns, A is secondary loop area, and ω is the frequency. The expected dependency of E_1 upon M has been introduced by defining $\kappa_e\equiv M\kappa_{e1}$, and the proportionality constant, S, accounts for the over-all sensitivity of the secondary detection system. S takes account of secondary tuning characteristics, primary-secondary coupling coefficient, detector-amplifier gain and relates the final output meter indication to the actual induced secondary emf. We assume h_{y0} proportional to $e^{-j\omega t}$ throughout the discussion.

On the other hand, if it is arranged that $M=0$ by carefully demagnetizing the sphere, and if the transmitter loop is rotated away from its position perpendicular to the receiver loop by a small angle, θ, the emf induced in the secondary can be derived from Eq. (3b) as

$$E_2=jS\theta nA\omega h_{y0}\mu_e. \tag{5}$$

It is assumed in (5) that all conditions of measurement are kept fixed as in (4) so that the constants may be eliminated. Defining $e_s\equiv E_1/M$ and $e_\theta\equiv E_2/\theta$, we have

$$e_s/e_\theta=\frac{|\kappa_{e1}|}{|\mu_e|}, \tag{6}$$

where magnitudes only are considered since the two measurements described by (4) and (5) are not related in phase.

Using Eq. (6) and values of $|\mu_e|$ as calculated from the independently determined permeability spectrum,[11] values of $|\kappa_{e1}|$ are obtained at each frequency. It is this quantity which has been chosen for comparison with theory, both because it is unique to domain rotation, as distinct from domain wall movement, and because it is a quantity which has already proven interesting in describing the ferromagnetic Faraday effect.

With the measurement procedure as outlined by Eq. (6), it is possible to reach frequencies at which ordinary loops cannot be used since all spurious effects which may occur can be made to cancel, provided the measurement is restricted to small values of M, and provided θ is so small that the conditions described by S are actually fixed throughout and can thus be eliminated. In the present case a single turn was used for both primary and secondary circuits, which were mounted directly between the poles of a small laboratory electromagnet. The secondary was wound tightly around the 5-mm diameter sphere of nickel ferrite and was cemented in place. The primary loop, diameter 1.0 cm, was mounted in a polystyrene holder and pivoted between the magnet poles. Its angle was adjustable by means of a 30-cm attached lever arm and micrometer screw. It is important to recognize that even though the loops can and do undergo self-resonance, and although primary-secondary capacitive coupling may be noticeable, the ratio procedure used nevertheless permits a measurement to be made. Inability to tune the loops affects the measurement only by reducing sensitivity.

The remainder of the apparatus is straightforward and is indicated schematically in Fig. 1(a). It was found necessary to include provision for tuning the magnet to 60-cycle series resonance in order to demagnetize the specimen thoroughly. Initial attempts by successive dc current reversals were unsuccessful, since it was difficult to avoid a small erratic residual magnetization from interfering in the application of Eq. (3b).

B. Operation

With the arrangement described, it has been possible to observe quite sizable signals. For instance employing 20 milliwatts of rf, the minimum leakage signal is about 1 μv (crystal output at 1 kc/sec modulation frequency) with the sphere demagnetized. Application of a steady field, sufficient to cause appreciable saturation, then causes the detector output to rise as high as a millivolt. It has, in addition, been possible to make measurements with steady fields as low as 5 oersteds external to the sphere indicating that the induction device can yield sensitivities adequate to permit study of effects in the initial permeability region.

The qualitative behavior of the observed signal for various loop angle adjustments is particularly significant in the present experiment, since it yields direct evidence of rotation independent of theoretical comparisons. When a steady field is applied and the comparisons. When a steady field is applied and the output signal rises above its minimum leakage value, it has been found impossible to reduce the signal to its

original value by readjusting the planes of the loops. The changing rf field responsible for the induced signal must therefore have components in both time and space quadrature, i.e., be rotational; if the inducing field were plane polarized, it could still be reduced to a minimum by proper loop-angle readjustment. The behavior thus cannot be explained by assuming a change in minimum leakage conditions, and the rotating rf field is thus probably attributable to the precessional nature of the polarized ferrite. In making measurements we have further found that changes in sample permeability resulting from partial saturation do not alter the tuning characteristics of the loops. Both primary and secondary have been broadly enough tuned so that the small changes in permeability do not appreciably change tuning conditions.

In making use of Eq. (6) to determine the ratio $|\kappa_e|/\mu_e|$ and thus $|\kappa_e|$ as a function of frequency,[11] the following procedure should apply under ideal operating conditions. With the loops adjusted to such an angle that a null signal results, the dc field is applied and the output indication per unit field, e_s, recorded. Then, with the sphere completely demagnetized and all other conditions unchanged, e_θ is measured by recording the output signal per unit angle as the primary loop is moved away from its normal position. Equation (6) may then be applied.

However, it has been found impossible in practice to reduce the minimum leakage signal to an absolute null. Experience has been similar to that described[12] in the case of nuclear induction in that voltages induced by currents flowing in surrounding conductors (pole faces, shields, etc.) may have a component out of phase with that resulting from the desired loop pick-up. In the present case, however, the signal from the ferrite is comparable with the leakage so that the total output, including both leakage and true signal, is not a linear function of the true signal itself. In order to obtain an output proportional to the voltage induced by the ferrite, a vector subtraction has been performed to eliminate the leakage voltage. This can be seen in Fig. 2(a) where the leakage voltage resulting after adjustment of the planes of the loops is shown as the ac vector e_l. Here e_l represents the minimum leakage obtainable for zero M by adjusting the relative loop angle. In practice it has been found that the total output signal becomes larger than e_l when M is increased from zero in one direction, whereas in the opposite direction the output at first decreases and then rises as before. In order to account for this behavior, it has been assumed that e_l remains fixed for all values of M encountered in the measurement, and that the output signal is composed of e_l plus a signal due to the precession of M. In Fig. 2(a) the situation is represented by showing the precession signal as e_s along the vertical axis and the leakage, e_l, at an arbitrary phase angle. The observed behavior can then be understood if it is noted that e_s is along the positive vertical axis for values of M of one direction but must be represented as e_s' along the negative vertical axis, i.e., be 180° out of phase with e_s, for values of M in the other. Since e_l has been assumed fixed, for one polarity of M the total output initially increases as shown by e_t, whereas for reversed values of M (and induced signals e_s') the total output at first decreases. The minimum value of total signal mentioned above can thus be seen to occur when the signal vector is at right angles to e_s and e_s'. This is indicated as e_m in the figure. Since it is the precession signal, e_s, which is desired for comparison with the theory of domain rotation, the correction for the contribution of e_l must be made as follows:

$$e_s=(e_t^2-e_m^2)^{\frac{1}{2}}-(e_l^2-e_m^2)^{\frac{1}{2}}. \tag{7}$$

In this way values of e_s have been obtained for use in Eq. (6).

The procedure for correcting e_θ is somewhat similar to that just outlined. In this case M is kept zero for all loop angles, and the minimum leakage can be represented as e_n of Fig. 2(b). Again, if it is assumed that the leakage signal is independent of loop angle for small changes in angle, the observed output signal, shown as e_t', can be assumed to be comprised of e_n plus the desired signal e_θ. The correction for leakage is therefore

$$e_\theta=[(e_t')^2-e_n^2]^{\frac{1}{2}}. \tag{8}$$

It is noted that the leakage vector e_l of Fig. 2(a) is identical with e_n of Fig. 2(b) although otherwise the two diagrams are independent.

In practice the use of the correction formulas (7) and (8) has led to values of e_s and e_θ which are proportional to the steady field, H, and to the loop angle (for small angles) respectively. This proportionality has been taken as evidence to justify the assumed leakage corrections.

Before describing the results obtained at various frequencies, it is of interest to consider qualitatively the output signal from the apparatus when the ferrite is not close to the demagnetized state, as above, but approaches magnetic saturation. As indicated in Fig. 1(a), an x-y chart recorder has been used during part of the work to aid in making adjustments. The arrange-

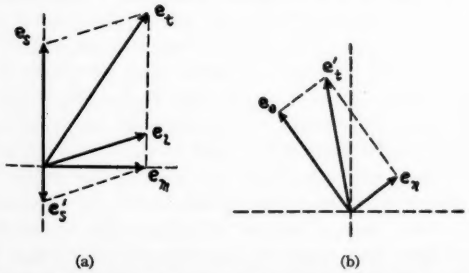


FIG. 2. Ac vector diagrams showing correction for effects of leakage signal. In (a) leakage is indicated by e_l, the desired precession signal by e_s, and the total observed signal by e_t. In (b) e_n is the residual leakage, e_θ is the signal induced by loop angle deflection, and their result is shown as e_t'.

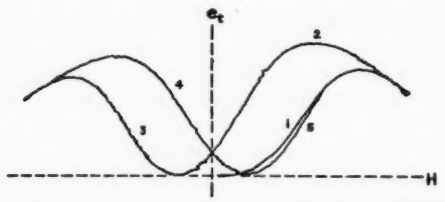


FIG. 3. Tracing from original x-y recorder chart showing total output of magnetic induction apparatus for various applied steady fields extending into region of saturation. Starting from demagnetized state (origin) the branches were numbered in sequence as recorded.

ment makes it possible to plot the output signal, designated as e_t in Fig. 2(a), for various magnet currents.

Starting at the origin in Fig. 3 (demagnetized condition) the output at first increases (curve 1) with applied field in a manner similar to the normal magnetization curve. However, when the external field approaches saturation, the effect of increasing net moment and attendant augmented signal is overcome by the shift in ferrimagnetic resonance frequency which accompanies the applied field. The latter effect tends to remove the dispersion region from the proximity of the measurement frequency toward the microwave region causing the precession of M at 50 Mc/sec to be considerably reduced. The maximum in branch No. 1 of Fig. 3 thus represents a balance between the two effects. Returning to lower applied fields along branch No. 2, a typical hysteresis pattern is apparently superimposed on the resonant frequency shift just mentioned. Back at zero applied field, the remanent magnetization (associated with the spherical sample as opposed to that of the infinite medium) yields a very noticeable signal as shown. Following branch No. 2 for reversed fields, the output drops to a value close to that associated with the initial completely demagnetized state. This occurs at the coercive field. The parabolic shape of the output signal around this point results from the square law response of the crystal. Branch No. 3 is also positive, since magnitudes only are observed. The rest of the hysteresis loop, branches No. 4 and 5, are similar to No. 1 and 2 with the exception that branch No. 4 does not rise as high as No. 2. We attribute this to the effect of leakage which may add to signals of one sense but subtract from those of the other.

C. Results

The qualitative results of this work have already become apparent from the description of operation. Essentially they are that an observable signal is induced in the secondary loop when a state of steady polarization is produced in the sample. When corrected for the effects of leakage, this signal has been found to be proportional to the applied field for fields such that the ferrite remains in the region of initial permeability. Under conditions of partial magnetic saturation, it has been found impossible to restore the minimum leakage signal achieved under demagnetized conditions. This behavior is consistent with the assumption of a precessing net magnetic moment, and has been associated with the precession of elementary domains comprising the net moment. On the other hand it has been found difficult to picture a way in which domain walls, as generally conceived, may produce the observed effect. The argument is essentially that the large number and random distribution of domain walls, if they were to exist in a polycrystalline sample, would cause cancellation of gyromagnetic effects and prohibit the formation of a rotating magnetic moment, an essentially cooperative phenomenon. If it is also assumed that the only effect of applying a steady field to the sample is to move the position of domain walls without changing the direction of spins within the domains (no domain rotation), then the bulk material must respond isotropically to an applied rf field and no induction signal could be observed. The qualitative behavior of the apparatus thus points toward domain rotation in the initial permeability region and discourages the concept of conventional domain walls as the sole contributing mechanism.

Quantitatively we have attempted to relate the size of the induction signal obtained to the predictions of the theory of domain rotation as given in II. To do this we have multiplied the observed values of $|\kappa_{el}|/|\mu_e|$, as obtained from Eq. (7) for each frequency, by $|\mu_e|$ as obtained from a separately determined permeability spectrum of the material. Here $|\mu_e|$ is an effective permeability of the spherical sample obtained from the true permeability components, μ_a' and μ_a'' using the procedure of II. Thus $|\kappa_{el}|$, again an *effective* permeability component, is obtained for the spherical sample. It represents the x-component of B produced by a y-component of H *external* to the sphere. The procedure of II may again be employed to obtain $|\kappa_a|$ from $|\kappa_{el}|$, and this in turn may be compared with theory. $|\kappa_a|$ is the off-diagonal permeability component for an infinite magnetically saturated medium.

In Fig. 4 the values of $|\kappa_a|/M$ determined in this way have been plotted as a function of frequency, and it can be seen that they are approximately proportional to frequency at low frequencies and that a rather broad maximum is reached around 200 Mc/sec. At higher frequencies the extent of the internal precession, as measured by $|\kappa_a|/M$, appears to drop off slowly. The large errors indicated around the experimental points arise principally from uncertainties in magnetic field measurements, but are increased by the conversion from effective to true values of μ and κ, using the procedures of II, and by mechanical instabilities in the apparatus. In certain cases, a number of independent determinations have been made to test reproducibility and the average of such readings has been used. The points at 250, 400, and 600 Mc/sec were obtained this way, and the errors plotted represent the spread of individual measurements.

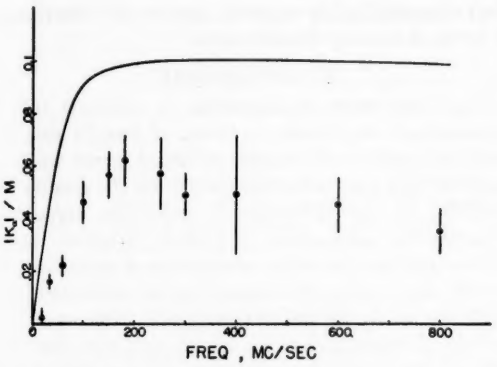


FIG. 4. Off-diagonal permeability component, $|\kappa_a|$, per unit magnetic moment *vs* frequency. Points are from present experiment on nickel ferrite, curve is from theory of domain rotation.

III. DISCUSSION

Also in Fig. 4 we have plotted theoretical values of $|\kappa_a|/M$ as obtained from Eq. (29) of II. The theory is based on an assumed frequency distribution of crystallites characterized by four parameters. Values of these parameters have been taken from the ordinary permeability spectrum of the particular nickel ferrite used in the experiment. Since the toroidal sample used to measure the magnetic spectrum and thus determine the parameters to be used in the theory had experienced similar but not identical processing to that experienced by the material used in our spherical sample, and since processing, particularly firing temperature, is very important in establishing the high-frequency characteristics of the material, the numbers chosen for insertion in the theory may be subject to error. They have, nonetheless, been used in order to compute the theoretical values shown.

In general, the experimental data agree qualitatively with prediction in that in both cases a distinct "knee" occurs at around 200 Mc/sec followed by a gradual decrease in $|\kappa_a|/M$ at higher frequencies. Although not shown in Fig. 4, the corresponding values of a μ'' as measured for the particular nickel ferrite here used also goes through a maximum close to this frequency; this is consistent with the domain rotation hypothesis in that the frequency at which the domain magnetic moments precess most strongly will also be the frequency at which the losses are greatest. This is clearly shown in Figs. 2 and 3 of II.

We cannot, however, attach any special significance to the factor of two discrepancy between theory and experiment. There are a number of ways in which it might be understood. As stated earlier, we have applied a correction factor giving $|\kappa_a|$ from the actually measured $|\kappa_{e1}|$ values. This amounts to a factor of the order of 20 which is quite sensitive to the choice of demagnetizing factor. In our case, we have assumed the rf field of the circular primary loop to be equivalent to a uniform field, since the sample diameter was less than half that of the primary. This assumption is only approximately valid and the resultant error might be included in the value of demagnetizing factor used to get $|\kappa_a|$ from $|\kappa_{e1}|$. Further, the theoretical curve represents a calculation which has been found to be most applicable in cases where the ferrite is sintered at a slightly lower temperature than the sample here used. There are therefore a number of considerations requiring further experimental study before any detailed quantitative comparison is possible. The qualitative results, however, clearly indicate the existence of domain rotation.

Possibly the most striking feature of the above results is the relatively low frequency at which this domain rotation in an equivalent internal anisotropy field appears to exist. This is striking because measurements on single crystals of nickel ferrite show the internal magneto-crystalline anisotropy field to be around 300 oersteds, corresponding to a natural ferrimagnetic resonance frequency of about 900 Mc/sec, using effective g-values and the fact that the bulk anisotropy field governs the precession of the two-sublattice ferrite.[15] Ordinarily one would expect the other contributions to the equivalent anisotropy of crystallites to *increase* this value rather than to decrease it making it hard to reconcile the present observations of a rotational resonance at 200 Mc/sec with other findings on single crystals.

In order to understand this discrepancy, it is important to distinguish clearly between the behavior of single crystals and that of the polycrystal we have studied. On a microcrystalline scale, a single crystallite may be subject to quite inhomogeneous demagnetizing fields, resulting from neighboring crystallites, of the order of the natural magneto-crystalline fields themselves. Thus the thickness of a domain wall, as ordinarily considered, may be quite different from that calculated for a single crystal. In fact the crystallites of our sample, being of the order of 1 to 3 microns in size, may well be of transition size between very small crystallites around a half micron where pronounced single domain behavior can be expected and larger sizes where a genuine multidomain situation exists. By using the procedure of Herring and Kittel,[16] the domain wall thickness in a large single crystal is found to be around a tenth micron. However, if an appreciable volume fraction of each crystallite contains spins directed away from directions of easy magnetization, as may be the case if the crystallites are of transition size, the magnetization of the crystallite as a whole may experience an average effective internal field much smaller than the 300 oersteds found in the very special case of a macroscopic single crystal. The dispersion region of such a polycrystal may thus occur at frequencies well below the expected 900 Mc/sec.

[15] F. Brown and D. Park, Phys. Rev. **93**, 381 (1954).
[16] C. Herring and C. Kittel, Phys. Rev. **81**, 869 (1951).

The hypothesis here considered thus removes the clear distinction between domain wall motion and domain rotation as the mechanism responsible for high frequency dispersion. It permits a large domain rotation effect to exist and cause natural ferrimagnetic resonance at relatively low frequencies as found here and elsewhere[1,6] and which is attributed to rotation of spins within regions of inhomogeneous magnetization. However, these inhomogeneous regions themselves may also be thought of as incipient domain walls which undergo resonant response in the applied rf field. If one accepts the above picture, it is possible to understand those cases where two distinct dispersion regions exist in the high-frequency magnetic spectrum of a ferrite[5] without introducing additional mechanism. Such materials may be thought of as those in which the polycrystalline sizes and magnetic parameters are such that genuine domain walls can form and respond in the manner usually visualized. However, in cases such as that studied here and elsewhere,[1,6] where only a single dispersion region may exist, conditions can be considered such that multidomain structures are only slightly approached, the largest contribution to initial permeability being from domain rotation. In this way it seems quite possible that the theory of domain rotation can achieve the observed degree of confirmation even when applied to magnetic spectra also describable in terms of moving domain walls.

IV. CONCLUSIONS

From this study it is possible to conclude that a phenomenon describable in terms of domain rotation definitely exists in the sample of nickel ferrite we have studied. This conclusion rests solely on the qualitative behavior of the ferrimagnetic induction apparatus. Quantitative application of domain rotation theory further yields a favorable comparison of gross features of the observations. In attempting to understand the existence of a natural ferrimagnetic resonance dispersion at frequencies as low as the presently observed 200 Mc/sec, it is suggested that a transition region between pure domain rotation and pure domain wall motions may exist in polycrystalline samples within which region a ferrite may exhibit aspects of both mechanisms. To this extent, the distinction between the two mechanisms becomes less physical than previously considered, and the high-frequency magnetic behavior of ferrites can be understood only after careful consideration of internal polycrystalline structure.

Note added in proof.—The present experiment cannot distinguish the situation in which domain wall motions throughout the sample are distributed in space and time quadrature so as to simulate the effects of domain rotation. Such a situation could occur, however, only if the walls were excited through coupling with domain rotation in agreement with the above conclusions.

Axial Ratios in Hexagonal Crystals*

J. W. McCLURE†‡
Department of Physics and Institute of Nuclear Studies, University of Chicago, Chicago, Illinois
(Received September 1, 1954; revised manuscript received January 12, 1955)

A theoretical study has been made of the deviations of the axial ratios of hexagonal metal crystals from the ideal close-packing value, both for pure metals and alloys. The deviation is expressed in terms of the "effective stress" which would tend to change the axial ratio in the close-packed configuration, and the observed elastic constants. Using a change of scale procedure, a formula for the "effective stress" is derived in terms of integrals over the electronic wave function in the close-packed configuration. When the Hartree-Fock approximation is used to evaluate the expression for the "effective stress," three contributions are found: (1) a "kinetic stress," (2) an "electrostatic stress," and (3) an "exchange stress." Estimates of each term for pure beryllium indicate that (1), which resembles the effects estimated by Jones and Goodenough, is most important; (2) is negligible; and (3) may be appreciable. It is found that the change in axial ratio with alloying is due to a different mechanism than that proposed by Jones. However, an argument is presented which leads to qualitatively the same conclusions as his concerning the connection between the change in axial ratio and the band structure.

1. INTRODUCTION

IF the forces between atoms in a hexagonal close-packed lattice were central forces between nearest neighbors only, the configuration for minimum energy would be that in which all nearest neighbor distances are equal.[1] For this ideal close-packed configuration the axial ratio; i.e., the ratio of the distance (c) between identical hexagonal planes to the distance (a) between nearest neighbors in these planes; would have the value $(8/3)^{\frac{1}{2}}=1.633$.

There are a number of metals with hexagonal crystal structures which have axial ratios that deviate from the ideal value by small amounts. A few of these metals are listed in Table I. The deviations from close-packing can be changed by changing the temperature and by alloying with another metal. Moreover, it is very interesting that the changes in axial ratio produced in a given metal by alloying with any other metal at constant temperature seem to depend chiefly upon the electron-to-atom ratio in the alloy.[2,3]

H. Jones has treated the problem of the change in axial ratio with alloying.[4] He made the hypothesis that the changes in axial ratio produced by alloying are due to changes in the Fermi energy arising from the distortion of the Brillouin zone. His theory makes use of the fact that the Brillouin zone for a hexagonal lattice is anisotropic and, as he considered cases in which the Fermi surface overlaps some of the zone faces, the Fermi surface is anisotropic. A change in axial ratio distorts the zone, forcing a distortion of the Fermi surface that, in general, changes the energy of the conduction electrons. Thus, in a manner of speaking, there exists an "effective stress" due to the conduction electrons, tending to change the axial ratio. The equilibrium axial ratio depends on a balance between the "effective electron stress" and the stress due to changes in all the other contributions to the total energy.

Jones assumed that upon alloying, the valence electrons of the solute atom go into the conduction bands of the solvent and increase the "effective stress" due to the conduction electrons. Thus he reasoned that the "effective electron stress," and thereby the axial ratio, is a function of the number of valence electrons per atom. He estimated the change in energy of each one-electron state with change in axial ratio by assuming that the energy of the state changes with shear like a free electron state with the same wave number. Under the above assumption, the change in axial ratio with alloying depends upon the band structure in the vicinity of the Fermi surface, as the structure determines the assignment of the extra valence electrons to positions in the Brillouin zone, and the assignment determines the change in "effective stress."

TABLE I. Summary of experimental data.

Metal	Beryllium	Magnesium	Zinc	Cadmium
c/a	1.565[a]	1.625[b]	1.856[c]	1.89[d]
ϵ_{obs}	−0.028	-0.003_3	0.091	0.107
μ (ev/atom)	15.8[e]	8.81[e]	6.45[e]	6.35[e]
σ (ev/atom)	−0.44	−0.028	0.56	0.66
η/μ	0.625	−0.047	−1.54	−1.52
$\partial\epsilon_{obs}/\partial x$[f]		0.023[b]	1.0[c]	

[a] P. Gordon, J. Appl. Phys. **20**, 908 (1949).
[b] See reference 3.
[c] E. A. Owen and L. Pickup, Proc. Roy. Soc. (London) **A140**, 179, 344 (1933); E. A. Owen and J. Iball, Phil. Mag. **17**, 433 (1934).
[d] F. Seitz, *The Modern Theory of Solids* (McGraw-Hill Book Company, Inc., New York, 1940), p. 6.
[e] The elastic constants of beryllium are those of L. Gold, Phys. Rev. **77**, 390 (1950). The other elastic constants are from the measurements of Wright and Grüneisen and Goens, as quoted in Gold's paper.
[f] The quantity x is the number of valence electrons per atom in the alloys.

* Based on a thesis submitted to the Department of Physics, University of Chicago, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.
† National Science Foundation Pre-doctoral Fellow.
‡ Present address: Department of Physics, University of Oregon, Eugene, Oregon.
[1] F. R. N. Nabarro and J. H. O. Varley, Proc. Cambridge Phil. Soc. **48**, 316 (1952), Part 2.
[2] G. V. Raynor, Proc. Roy. Soc. (London) **A174**, 457 (1940).
[3] W. Hume-Rothery and G. V. Raynor, Proc. Roy. Soc. (London) **A177**, 27 (1940); G. V. Raynor, Proc. Roy. Soc. (London) **A180**, 451 (1942); R. S. Busk, J. Metals **188**, Trans. 1460 (1950); R. S. Busk, J. Metals **4**, Trans. 207 (1952).
[4] H. Jones, Proc. Roy. Soc. (London) **A147**, 400 (1934); Phil. Mag. **41**, 633 (1950).

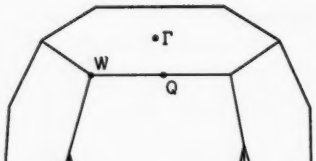


FIG. 1. The Brillouin zone for the hexagonal close-packed lattice which contains two electrons per atom.

In the 1950 work, Jones discussed the changes in axial ratio of magnesium induced by alloying. He attempted to establish the band structure of magnesium by constructing a simple model to fit the experimental data. In order to explain Jones's model, we refer to Fig. 1 which shows the second Brillouin zone for the hexagonal close-packed lattice. This zone contains exactly two electrons per atom. All of the valence electrons of pure magnesium would completely fill the zone; but magnesium is a metal and we know that in the ground state there is some overlapping into higher zones and some holes (unoccupied states) in the zone of Fig. 1. In analogy to the nearly free electron model, Jones assumed that the energy increases going away from the origin in reciprocal space. He took points K and Γ as possible overlap sites and points Q as possible hole sites.

Under the assumptions of the Jones theory, an electron added at point Γ tends to increase the axial ratio. (The energy of an electron at point Γ is decreased if the zone is distorted so that Γ moves nearer the origin. Such a distortion corresponds to an increase of the axial ratio of the direct lattice.) In like manner, an electron added at point K tends to decrease the axial ratio, and one added at point Q tends to increase the axial ratio. In order to fit the experimental data, Jones assumed that in pure magnesium there are overlaps at K and holes at Q, and that overlap takes place at Γ for valence electron to atom ratios slightly larger than two (2.0075).

The assumption that overlap occurs at Γ for a valence-electron-to-atom ratio of 2.0075 was made because there is experimental evidence that the derivative of the axial ratio with respect to the valence electron to atom ratio shows a sharp change at this point. The idea that the sharp change is due to the onset of overlap at the point Γ was proposed by Raynor.[2]

Such an onset of overlap should be noticeable in other properties of magnesium alloys. In particular, Schindler and Salkovitz[5] have investigated the change in the Hall constant in magnesium alloys. They find evidence of a sharp change in the slope of the Hall constant *versus* the valence-electron-to-atom ratio curve in the neighborhood of 2.013 valence electrons per atom.

Goodenough[6] has applied a similar theory to the case in which the Fermi surface is near to, but does not touch a set of Brillouin zone faces. Using the nearly-free electron model as a starting point, he argues that there is an effective attraction between the Fermi surface and the zone faces; i.e., due to the depression of energy beneath a zone face, the total energy of the electrons could be lowered by moving the zone face nearer to the origin in reciprocal space.

We illustrate Goodenough's ideas with the following example. Consider a pair of parallel zone faces (represented by P in Fig. 2) which are moved closer to the origin (P') as the result of a deformation. Figure 2 represents the variation of the electron energy as a function of wave number along a line perpendicular to the two planes and passing through the origin. The energy before deformation is represented by the solid curve and after deformation by the dashed curve. It is seen that the energies of the electrons in positions corresponding to F are lowered as a result of the deformation. Thus if the Fermi surface does not touch any of the zone faces, the total electronic energy is lowered.[7] The change in energy with change in axial ratio given by the Goodenough mechanism gives rise to an "effective stress" which determines the axial ratio as in the theory of Jones.

The attractive feature of the above theories is that they seem to afford a way of utilizing the experimentally observed changes in axial ratio with alloying to obtain information about the band structure in the neighborhood of the Fermi surface. Such information, which is very interesting as it determines other properties of the metal, is difficult to obtain by fundamental calculation. Thus, any method which holds promise of easily obtainable information about the band structure deserves careful investigation. Both theories have a certain plausibility, but they both fail to take account of a number of contributions to the "effective stress" which one would expect to be important. For example, the assumption that the change in "effective stress" with alloying is due to the change in the stress arising from the "Fermi energy" alone, neglects contributions to the "effective stress" coming from changes in the energy of the bottom of the band, Coulomb, exchange, and correlation energies, and energy of lattice vibrations, all of which contribute to the total energy of the crystal. Further, the application of the nearly free electron model is equivalent to ignoring two important facts: the potential in which an electron inside a metal moves is very strong, and the valence electron wave functions

[5] A. I. Schindler and E. I. Salkovitz, Phys. Rev. **91**, 1320 (1953).

[6] J. B. Goodenough, Phys. Rev. **89**, 282 (1953).

[7] If the Fermi surface does touch some of the zone faces, some electrons must be displaced by the movement of the zone faces to other points on the Fermi surface, so that their energies are increased. Whether the total energy is increased or decreased will depend on the details of the band structure and the Fermi surface.

must be orthogonal to the closed-shell electron wave functions. Thus, since the behavior of electrons in real metals differs appreciably from their behavior in a free electron gas, it is not to be expected that results based on the nearly free electron model are valid.

In treating alloys, Jones assumed that the one-electron wave functions and energy spectrum of a dilute alloy are identical to those of the pure metal. Under such an assumption the effects due to alloying come from the increased (or decreased) number of valence electrons in the Fermi distribution. However, it is to be expected that there are other effects which are just as important; e.g., the interaction of the electrons already present with the impurity atoms.

In the present work, we develop a deductive theory of the deviations of the axial ratio from close-packing, which starts from fundamentals and applies to both pure metals and alloys. We display the factors important in determining the axial ratio in order to elucidate the basis of the theories of Jones and Goodenough. We carry the theory far enough to make numerical estimates of the divers factors in the case of pure beryllium, and to investigate the change in axial ratio with alloying.

In Sec. 2 the "effective stress" is defined, and its role in determining the axial ratio is discussed. A general expression for the "effective electron stress" is derived in Sec. 3. In Sec. 4 an expression for the "effective stress" for a pure metal is obtained using the Hartree-Fock approximation. Section 5 contains a discussion of the "effective stress" in the Wigner-Seitz-Slater approximation. In Sec. 6, the relation of the present work to that of Jones and Goodenough is discussed. A numerical application to pure beryllium is presented in Sec. 7. The change of lattice parameters with alloying is discussed in Sec. 8. Finally, in Sec. 9, the results of the paper are summarized and conclusions drawn.

2. ROLE OF THE EFFECTIVE STRESS

We now derive an expression for the value of the axial ratio at absolute zero of temperature, in terms of the observed elastic constants and the "effective stress" tending to shear the crystal. In this section we also briefly discuss the question of the temperature dependence of the axial ratio.

The equilibrium configuration of a crystal at absolute zero of temperature and under no external stress is that configuration for which the total energy is a minimum. We will expand the energy of the crystal in a Taylor's series about a configuration (hereafter called the ideal configuration) which has the observed atomic volume and the ideal axial ratio corresponding to perfect close-packing. We expand in terms of a parameter ϵ, defined by the equations

$$c=c_0(1+\epsilon); \quad a=a_0(1-\tfrac{1}{2}\epsilon), \tag{2.1}$$

where c_0 and a_0 are the values of c and a in the ideal configuration. It is seen that, to first order in ϵ, a change in ϵ represents a change in axial ratio at constant volume. To first order we have

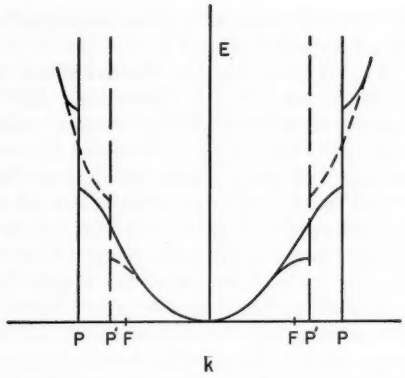


FIG. 2. Schematic representation of the variation of the electron energy with wave number for two configurations of the lattice.

$$c/a=(8/3)^{\frac{1}{2}}(1+\tfrac{3}{2}\epsilon). \tag{2.2}$$

Neglecting terms higher than second order, we write for the energy at absolute zero

$$E=E_0-\sigma\epsilon+\tfrac{1}{2}\mu\epsilon^2, \tag{2.3}$$

where E_0 is the energy in the ideal configuration and σ and μ are constant coefficients. Minimizing expression (2.3) with respect to ϵ, we find

$$\epsilon_{\min}=\sigma/\mu. \tag{2.4}$$

A theoretical calculation of σ and μ would, of course, give a theoretical value of $\epsilon_{\min}$ and the axial ratio. However, we shall confine ourselves to a theoretical calculation of σ, the "effective stress," as this quantity is of most physical interest. We obtain an approximate value of μ using the observed elastic constants. The quantity μ is given by the second derivative (at constant volume) of E with respect to ϵ, evaluated in the ideal configuration. To the extent that the quadratic approximation is valid [Eq. (2.3)], the value of the derivative is the same in the ideal configuration as in the observed configuration. But the latter is a combination of the elastic constants observed at absolute zero. The expression for μ is given by Jones in the 1950 paper.

In Table I are listed the observed values of $\epsilon_{\min}$ and the values of σ necessary to produce the observed deviations from close-packing. Though the data in the table was not obtained at absolute zero, the differences of the values listed from the correct values at absolute zero are probably small. For example, the order of magnitude of thermal expansion coefficients is 10^{-5}, so that up to 1000°K the change in $\epsilon_{\min}$ is of the order of 0.01. Though the elastic constants used to compute μ were not measured as a function of temperature, the adiabatic compressibility and rigidity modulus of beryllium show very little dependence on temperature.[8] The fact

[8] W. C. Overton, J. Chem. Phys. 18, 113 (1950). The apparent discontinuity of the rigidity modulus in Table II is the result of a misprinted decimal point in the values for 140°K and below.

that the observed values of $\epsilon_{\min}$ are small justifies the approximation made in Eq. (2.3).

Thus, a calculation of σ, the "effective stress," serves to calculate the axial ratio. It follows from (2.3) that σ is the negative of the derivative (at constant volume) of the energy with respect to ϵ, evaluated in the ideal configuration. The total energy at absolute zero is equal to the total energy with the nuclei held fixed on the sites of a perfect lattice (which we shall call the electronic energy) plus the zero point energy of the nuclear vibrations. As the latter energy is much smaller than the former, we shall neglect it in this paper.[9] Thus, in the present approximation, the axial ratio is determined by the negative of the derivative of the electronic energy, which we shall call the "effective electron stress."

We shall close this section with a few brief comments on the temperature dependence of the axial ratio. To determine the axial ratio at a finite temperature, one must minimize the Helmholtz free energy with respect to the axial ratio and atomic volume. The second order coefficients in the Taylor's expansion of the free energy can be approximated by combinations of the isothermal elastic constants. The "effective stress" tending to change the axial ratio or the volume is found by taking the appropriate first derivative of the free energy.

Considering first the contribution of the lattice vibrations to the free energy, we note that the derivative with respect to ϵ (or the volume) will depend on temperature in a manner resembling the total vibrational energy of the lattice. If the derivative is assumed to be proportional to the vibrational energy, the theory of Grüneisen is obtained.[10]

The derivative of the electronic energy with respect to ϵ (or the volume) likewise behaves qualitatively like the total electronic energy. Provided that there is no peculiarity in the density of states at the Fermi level, the "effective electron stress" is equal to the stress at absolute zero plus a term proportional to the square of the absolute temperature. The thermal expansion coefficients would then contain terms proportional to the absolute temperature. The possibility of the existence of such a linear term in the volume expansion coefficient was pointed out by Visvanathan.[11] The linear contribution to the thermal expansion coefficient of the axial ratio may be too small to measure, however.

3. GENERAL EXPRESSION FOR THE EFFECTIVE ELECTRON STRESS

In this section we derive an exact expression for the "effective electron stress" of a single state of the system in terms of integrals over the wave function of the ideal configuration. The expression derived is in the form of a generalized virial theorem. By generalizing the procedure still further, it would be possible to calculate the "effective stress" due to the lattice vibrations, though we have not so attempted in this paper.

We imagine that initially the nuclei are held rigid in the ideal configuration. We will compute the first-order change in energy in going to a slightly different configuration (hereafter called the strained configuration) derived from the ideal configuration by application of the infinitesimal strain tensor $\mathbf{s}$. We have

$$\mathbf{R}_\mu' = (1+\mathbf{s})\cdot\mathbf{R}_\mu, \tag{3.1}$$

where $\mathbf{R}_\mu$ and $\mathbf{R}_\mu'$ are the positions of the μth nucleus in the ideal and strained configurations, respectively. In the special case of the change in axial ratio, the tensor $\mathbf{s}$ is given in terms of the scalar ϵ defined in (2.1) by

$$s_{xx}=s_{yy}=-\tfrac{1}{2}\epsilon; \quad s_{zz}=\epsilon, \tag{3.2}$$

all other components being zero.

We write the Hamiltonian for the ideal configuration:

$$H_0=\textstyle\sum_j \mathbf{p}_j^2/2m+V(\mathbf{R}_\mu,\mathbf{r}_j), \tag{3.3}$$

where the sum j is over all electrons, $\mathbf{r}_j$ and $\mathbf{p}_j$ are the position and momentum vectors of the jth electron, m is the mass of the electron and the potential V is the electrostatic potential energy of all the electrons and nuclei. We have neglected all other contributions to the energy as, for example, spin-orbit coupling. The Hamiltonian for the strained configuration is then

$$H_1=\textstyle\sum_j \mathbf{p}_j^2/2m+V[(1+\mathbf{s})\cdot\mathbf{R}_\mu, \mathbf{r}_j]. \tag{3.4}$$

We now make a change of variables such that

$$\mathbf{r}_j=(1+\mathbf{s})\cdot\mathbf{r}_j'. \tag{3.5}$$

Inserting (3.5) into (3.4) and expanding to first order in $\mathbf{s}$ yields

$$H_1(\mathbf{r}_j')=H_0(\mathbf{R}_\mu,\mathbf{r}_j')-\textstyle\sum_j \mathbf{p}_j'\cdot\mathbf{s}\cdot\mathbf{p}_j'/m \\ +\sum_j \mathbf{r}_j'\cdot\mathbf{s}\cdot\nabla_j' V+\sum_\mu \mathbf{R}_\mu\cdot\mathbf{s}\cdot\nabla_\mu V, \tag{3.6}$$

where $\mathbf{p}_j'$ denotes the momentum conjugate to $\mathbf{r}_j'$.

The advantage of making the change in variables is that the eigenfunctions of (3.6), expressed as functions of the $\mathbf{r}_j'$, satisfy the same boundary conditions as the eigenfunctions of (3.3), since the $\mathbf{r}_j$ run over a cell in the strained lattice as the $\mathbf{r}_j'$ run over a cell in the ideal lattice. Thus we may find the eigenvalue of (3.6) by standard perturbation theory.

Because we are interested in the first-order change in energy, we need only use first-order perturbation theory; thus we find for the first-order change in energy:

$$\delta E(\mathbf{s})=\langle-\textstyle\sum_j \mathbf{p}_j\cdot\mathbf{s}\cdot\mathbf{p}_j/m \\ +\sum_j \mathbf{r}_j\cdot\mathbf{s}\cdot\nabla_j V+\sum_\mu \mathbf{R}_\mu\cdot\mathbf{s}\cdot\nabla V\rangle, \tag{3.7}$$

[9] It is possible that the derivative of the zero-point energy with respect to ϵ could be of the order of magnitude of the zero-point energy itself. As the order of magnitude of the zero-point energy is 0.1 ev per atom, the resultant stress would not be negligible, though it still would be less important than the "electronic stress," with the possible exception of magnesium.

[10] For a discussion of the Grüneisen theory, together with a review of the data on thermal expansion of anisotropic metals, see B. G. Childs, Revs. Modern Phys. **25**, 665 (1953).

[11] S. Visvanathan, Phys. Rev. **81**, 626 (1951).

where the angular brackets denote the expectation value taken with the wave function for the ideal configuration and the primes have been dropped from the variables of integration.

Equation (3.7) is exact in the case that the expectation value is taken with the exact wave function and is seen to provide a generalized virial theorem.[12]

An approximate value of the "effective stress" can be obtained by evaluating (3.7) with an approximate wave function. The approximate "effective stress" calculated in such a fashion is not necessarily equal to the derivative of the energy calculated in the same approximation. (I.e., the approximate derivative of the energy may not be equal to the derivative of the approximate energy.) However, the two ways of calculating the approximate stress do give the same answer in the important case of the Hartree-Fock approximation. The equivalent result to (3.7) in the Hartree-Fock approximation was derived by Fuchs and Peng.[13]

We now note that the expression (3.7) can be rewritten as

$$\delta E(\mathbf{s}) = C + B, \tag{3.8a}$$

$$C = -(i/\hbar)\sum_j \langle [H_0, \mathbf{r}_j \cdot \mathbf{s} \cdot \mathbf{p}_j] \rangle, \tag{3.8b}$$

$$B = \sum_\mu \mathbf{R}_\mu \cdot \mathbf{s} \cdot \langle \nabla_\mu V \rangle, \tag{3.8c}$$

where the square brackets in (3.8b) denote the commutator. In evaluating expressions (3.8) we have the choice of finding the total value for a finite sample, or the value per unit volume in a very large sample. We shall use the latter procedure in treating pure metals, and the former in discussing alloys. We now use the first procedure to discuss the physical meaning of the expressions (3.8).

If the sample is finite, the integrations in (3.8) may be taken over all space. We first consider (3.8b); carrying out a partial integration on the variable $\mathbf{r}_j$ we obtain

$$C = -(i/\hbar)\sum_j \int d^3r_1 \cdots \times \int d^3r_N [(H_0\Psi)^* \mathbf{r}_j \cdot \mathbf{s} \cdot \mathbf{p}_j \Psi - \Psi^* \mathbf{r}_j \cdot \mathbf{s} \cdot \mathbf{p}_j H_0 \Psi] + (\hbar/2mi)\sum_j \int d^3r_1 \cdots \int d\mathbf{S}_j \cdots \times \int d^3r_N [(\nabla_j \Psi)^* \mathbf{r}_j \cdot \mathbf{s} \cdot \mathbf{p}_j \Psi - \Psi^* \nabla_j (\mathbf{r}_j \cdot \mathbf{s} \cdot \mathbf{p}_j)\Psi], \tag{3.9}$$

where Ψ is (in this discussion) the exact many-electron wave function. The first term in (3.9) is zero as Ψ is an eigenfunction of H_0 and the second is zero as the wave function becomes vanishingly small at infinity.

Thus, in the case under discussion, the entire contribution to δE comes from B. By Feynman's theorem,[14] the quantity $\langle \nabla_\mu V \rangle$ is equal to the negative of the force on the μth nucleus due to all the other particles in the metal. As the total force on each nucleus must be zero in order to maintain the configuration, an outside force equal to $\langle \nabla_\mu V \rangle$ must be applied to the μth nucleus. The quantity B is then seen to be equal to the work done by the outside forces during the deformation described by $\mathbf{s}$.

As deep in the crystal each atom is surrounded symmetrically by its neighbors, $\langle \nabla_\mu V \rangle$ is appreciable only for the atoms near the surface. In order to calculate δE by the present method it would be necessary to know the variation of the potential near the physical surface.[15] Such a calculation would be quite difficult, though we shall see in Sec. 8 that the difficulty can be obviated if one is interested only in the change in δE with alloying.

4. EFFECTIVE ELECTRON STRESS FOR A PURE METAL IN HARTREE-FOCK APPROXIMATION

In this section an expression for the "effective stress" per unit volume is derived using the Hartree-Fock approximation. We imagine that the sample is infinitely large, and that the wave function is normalized inside a large volume Ω. We calculate δE inside a smaller volume ω, divide by ω and take the limit as both ω and Ω become infinite.

In carrying out the above limiting procedure, care must be taken to avoid divergences introduced by the fact that the coordinates appear linearly in the expression for δE [Eqs. (3.8)]. As the total potential is a function of differences of coordinates, we could have made the Taylor's expansion [Eq. (3.6)] in terms of differences of coordinates. Such an expansion in differences of coordinates would converge like the potential itself. We could obtain the expression (3.7) from the expression involving differences of coordinates by a change of variables, provided that the regions of integration (or summation) of all coordinates are the same. Thus, in calculating δE inside ω, we shall include only interactions between particles inside ω, which guarantees that the correct limiting value of δE will be obtained.

In order to facilitate the calculation outlined above, it is convenient to introduce a quantity $V_o(\mathbf{r}_j)$, which is the instantaneous potential of interaction between the jth electron and all the particles outside ω. With the use of $V_o(\mathbf{r}_j)$, the contribution of the volume ω to C can now

[12] A tensor virial theorem has been discussed by E. N. Parker [Phys. Rev. **94**, 1441 (1954)] in another connection.

[13] K. Fuchs and H. W. Peng, Proc. Roy. Soc. (London) **A180**, 451 (1942). These authors used the change of coordinates procedure and did a perturbation calculation on the one-electron Hartree-Fock wave functions. They also derived the surface integral form [Eq. (4.10) of the present paper], but expressed the remaining terms in a different form than in the present paper. Their derivation applies only to cases in which the wave functions are periodic, while our results can be applied to alloys.

[14] R. P. Feynman, Phys. Rev. **56**, 340 (1939).

[15] A calculation by this method would be similar to the calculation of the work function. See, for example, J. Bardeen, Phys. Rev. **49**, 653 (1936) or R. Smoluchowski, Phys. Rev. **60**, 661 (1941).

be written

$$C(\omega)=-(i/\hbar)\sum_j\int_\Omega d^3r_1\cdots\int_\omega d^3r_j\cdots$$
$$\times\int_\Omega d^3r_N\Psi^*[H_0-V_o,\mathbf{r}_j\cdot\mathbf{s}\cdot\mathbf{p}_j]\Psi, \quad (4.1)$$

where Ψ is the many-electron wave function.

We now make use of the Hartree-Fock approximation and let the wave function Ψ be a normalized determinant of one-electron spin orbitals $\psi_\alpha(\mathbf{r}_j)$ which satisfy the Hartree-Fock equation $H_F\psi_\alpha=\mathcal{E}_\alpha\psi_\alpha$, where H_F is the one-electron Hartree-Fock Hamiltonian and $\mathcal{E}_\alpha$ is the one-electron Hartree-Fock eigenvalue. We make the convention that the quantum number α stands for both spatial and spin quantum numbers and that when $\mathbf{r}_j$ appears as the argument of a ψ_α it stands for both spatial and spin variables, integration over $\mathbf{r}_j$ being understood to include summation over spin variables. We normalize the ψ_α to unity in the large volume Ω.

Consider first the part of (4.1) which depends upon the full Hamiltonian. We perform a partial integration on the variable $\mathbf{r}_j$, obtaining an expression identical to (3.9) except that the regions of integration are different. When the determinantal wave function is substituted into the expression, the term analogous to the first term on the right in (3.9) vanishes, as integration over all coordinates except $\mathbf{r}_j$ transforms the total Hamiltonian into the Hartree-Fock Hamiltonian. Because of the *ortho*-normality of the one-electron orbitals, the surface integral reduces to a sum of integrals involving one ψ_α at a time. The part of (4.1) which depends upon the commutator of V_o gives rise to terms representing an electrostatic interaction of the electrons inside ω with all particles outside of ω, and an exchange interaction between the electrons inside ω and the electrons outside of ω. The expression for C in terms of the one-electron orbitals is

$$C(\omega)=D+G+X, \quad (4.2)$$
$$D=(-i\hbar/2m)\sum_\alpha\int_\omega \mathbf{dS}\cdot[(\nabla\psi_\alpha)^*\mathbf{r}\cdot\mathbf{s}\cdot\mathbf{p}\psi_\alpha-\psi_\alpha^*\nabla(\mathbf{r}\cdot\mathbf{s}\cdot\mathbf{p})\psi_\alpha], \quad (4.2a)$$
$$G=-\int_\omega d^3r_1\rho(\mathbf{r}_1)\mathbf{r}_1\cdot\mathbf{s}\cdot\nabla_1\times\int_o d^3r_2\mathrm{P}(\mathbf{r}_2)/|\mathbf{r}_1-\mathbf{r}_2|, \quad (4.2b)$$
$$X=\sum_{\alpha\beta}\int_\omega d^3r_1\int_o d^3r_2\psi_\alpha^*(\mathbf{r}_1)\times\psi_\beta(\mathbf{r}_1)\psi_\beta^*(\mathbf{r}_2)\psi_\alpha(\mathbf{r}_2)\mathbf{r}_1\cdot\mathbf{s}\cdot\nabla_1 e^2/|\mathbf{r}_1-\mathbf{r}_2|, \quad (4.2c)$$

where the sum α is over occupied states and the subscript o indicates that the integration is over the volume outside of ω. The quantity ρ is the average charge density due to the electrons, and P is the total average charge density. They are given by

$$\rho(\mathbf{r})=-e\sum_\alpha|\psi_\alpha(\mathbf{r})|^2, \quad (4.3a)$$
$$\mathrm{P}(\mathbf{r})=\rho(\mathbf{r})+\sum_\mu eZ\delta(\mathbf{r}-\mathbf{R}_\mu), \quad (4.3b)$$

where $-e$ is the charge on an electron, Ze is the charge on a nucleus, and $\delta(\mathbf{r})$ is the Dirac delta function.

We next make use of the periodicity of the wave functions to transform G into a lattice sum inside ω. We then shall combine it with B to obtain a simple result. We first express G as a sum of integrals over atomic polyhedra. These polyhedra are the same as the Wigner-Seitz cells, the polyhedron about a given atom being the smallest volume contained by the perpendicular bisecting planes of the lines joining the atom in question to all other atoms. A polyhedron and a few of its neighbors are shown in Fig. 3. These polyhedra have the same symmetry as the space lattice; in particular, for the hexagonal close-packed lattice, the total charge and the dipole moment within a polyhedron are zero.

Denoting the polyhedron about the μth atom as $[\mu]$, we have for G,

$$G=-\sum_\mu^\omega\int_{[\mu]}d^3r_1\rho(\mathbf{r}_1)[\mathbf{R}_\mu+(\mathbf{r}_1-\mathbf{R}_\mu)]\cdot\mathbf{s}\cdot\nabla_\mu\times\int_o d^3r_2\frac{\mathrm{P}(\mathbf{r}_2)}{|\mathbf{r}_1-\mathbf{r}_2|}, \quad (4.4)$$

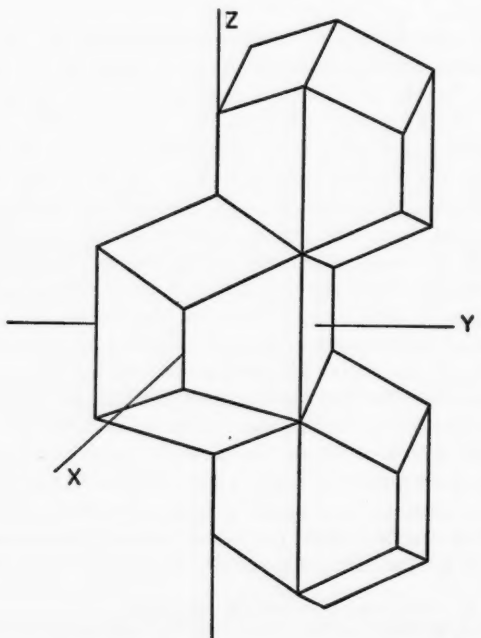


FIG. 3. The central atomic polyhedron and a few of its neighbors

where the superscript ω on the summation sign means that the summation is over those $\mathbf{R}_\mu$ inside ω. We specify that the surface of ω coincides with the bounding surfaces of atomic polyhedra. Thus, the field strength inside ω due to the charges outside ω will be appreciable only near the surface, there being no net charge outside ω or surface dipole moment on the surface of ω. Since, in (4.4), $\mathbf{r}_1-\mathbf{R}_\mu$ is always of the order of magnitude of an interatomic distance, the part of G proportional to $\mathbf{r}_1-\mathbf{R}_\mu$ will give a term proportional to the area of the surface, and may be neglected in the limit of infinite volume.

In the remaining part of (4.4), we write the integral over the region outside of ω as an integral over all space minus an integral over the volume ω,

$$G=-\sum_\mu^\omega \mathbf{R}_\mu\cdot\mathbf{s}\cdot\int_{[\mu]} d^3r_1\rho(\mathbf{r}_1)\nabla_1 \times\left[\int_\Omega d^3r_2\frac{\mathrm{P}(\mathbf{r}_2)}{|\mathbf{r}_1-\mathbf{r}_2|}-\int_\omega d^3r_2\frac{\mathrm{P}(\mathbf{r}_2)}{|\mathbf{r}_1-\mathbf{r}_2|}\right]. \quad (4.5)$$

The contribution of the first term in the square brackets to the integral over the polyhedron represents the average force on the electrons in the polyhedron, and vanishes by symmetry. Expressing the remaining integral over ω as a sum over polyhedra, we have

$$G=\sum_{\mu\nu}^\omega \mathbf{R}_\mu\cdot\mathbf{s}\cdot\int_{[\mu]} d^3r_1\int_{[\nu]} d^3r_2\rho(\mathbf{r}_1)\mathrm{P}(\mathbf{r}_2)\nabla_1\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}. \quad (4.6)$$

Remembering that we must count only interactions between particles inside ω, we express B [Eq. (3.8c)] as

$$B=\sum_{\mu\nu}^\omega Ze\mathbf{R}_\mu\cdot\mathbf{s}\cdot\nabla_\mu\int_{[\nu]} d^3r_2\frac{\mathrm{P}(\mathbf{r}_2)-Ze\delta(\mathbf{r}_2-\mathbf{R}_\mu)}{|\mathbf{r}_2-\mathbf{R}_\mu|}, \quad (4.7)$$

where the delta function cancels the interaction of the μth nucleus with itself. By summing these two expressions and averaging over an interchange of μ and ν, we find

$$S=B+G=\tfrac{1}{2}\sum_{\mu\nu}^{\omega\prime}(\mathbf{R}_\mu-\mathbf{R}_\nu)\cdot\mathbf{s} \cdot\int_{[\mu]} d^3r_1\int_{[\nu]} d^3r_2\mathrm{P}(\mathbf{r}_1)\mathrm{P}(\mathbf{r}_2)\nabla_1\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}. \quad (4.8)$$

We have finally expressed the "effective stress" due to a many-electron Hartree-Fock state as the sum of three terms,

$$\delta E(\mathbf{s})=D+S+X, \quad (4.9)$$

where the expressions for the individual terms are given by Eqs. (4.2a), (4.8), and (4.2c), respectively. We next shall discuss the physical interpretations of each of the three terms.

We begin by discussing term D [Eq. (4.2a)]. We make use of the periodicity of the wave functions to reduce D to an integral over the surface of a unit cell, obtaining

$$D(\omega)=\sum_\alpha D_\alpha=(\omega/\Delta)(-i\hbar/2m)\sum_\alpha\int_\Delta \mathbf{dS} \cdot[(\nabla\psi_\alpha)^*\mathbf{r}\cdot\mathbf{s}\cdot\mathbf{p}\psi_\alpha-\psi_\alpha^*\nabla(\mathbf{r}\cdot\mathbf{s}\cdot\mathbf{p})\psi_\alpha], \quad (4.10)$$

where D_α stands for the contribution to D from the one-electron state α. The symbol Δ stands for the unit cell so that the factor ω/Δ is the number of unit cells in ω. Due to the fact that in the evaluation of term S, the surface of ω was chosen to coincide with the surfaces of atomic polyhedra, the unit cell used in expression (4.10) must be composed of two neighboring atomic polyhedra, in adjacent planes (see Fig. 3). Because the contributions to D come from the region between atoms, the valence electrons will be most important in determining its value. In metals in which the interatomic distance is large compared to the ionic radius, the contribution from inner shell electrons should be negligible. In the other metals, however, the contributions from these electrons may have to be taken into account.

It is possible to construct a classical analog of term D. Consider a number of electrons per unit volume n having velocity $\mathbf{v}$ and momentum $\mathbf{p}$. The flux of particles through the surface element $\mathbf{dS}$ is $n\mathbf{dS}\cdot\mathbf{v}$, and the force on the surface element due to the momentum transport is $n(\mathbf{dS}\cdot\mathbf{v})\mathbf{p}$. Now suppose that the cell is deformed by a strain $\mathbf{s}$. The work done by this force acting through the displacement $\mathbf{s}\cdot\mathbf{r}$ is given by

$$\delta E=-n\int_S(\mathbf{dS}\cdot\mathbf{v})(\mathbf{r}\cdot\mathbf{s}\cdot\mathbf{p}). \quad (4.11)$$

Equation (4.11) can be seen to be a classical analog to (4.10) with $\mathbf{v}$ taking the place of the current operator. Thus we call term D the "kinetic stress" due to the electrons. We shall see in Sec. 6 that the term D corresponds to the stresses estimated by Jones and Goodenough.

Now we consider term S. Inspection of the structure of Eq. (4.8) shows that S represents the work done against the electrostatic forces between polyhedra in carrying out the deformation described by $\mathbf{s}$. As each polyhedron is electrically neutral, it is to be expected that S is small and that is what we find in Sec. 7.

The term X represents an exchange effect. It arises from the increased attraction of the medium outside of ω for an electron in ω, due to the fact that electrons of the same spin are kept apart by the determinantal wave function. A crude estimate of X in Sec. 7 indicates that it may be appreciable.

Finally, we note that the formalism developed in this section can be applied to find the linear change in energy of any solid under deformation. Note that it gives the change in energy of the total solid, not the change in one-electron energy parameters. To find the exact

change in Hartree-Fock energy parameters, a self-consistent calculation would be necessary. Such a calculation has been discussed by Bardeen.[16] However, the self-consistent calculation can be avoided in cases in which Koopmans's theorem[17] is valid. For example, the change in a certain one-electron energy parameter can be calculated by calculating the change in total energy for a state in which the one-electron state is occupied, and subtracting the change in energy of another state which differs from the first in that the particular one-electron state is empty, and one electron is removed from the crystal. The change with strain of the energy parameters is of interest in connection with the change with strain in the energy gap of a semiconductor,[18] the explanation of the elastoresistance of germanium,[19] and the deformation potential theory.[20]

5. EFFECTIVE STRESS IN THE WIGNER-SEITZ-SLATER APPROXIMATION

In treating the cohesive energy of monovalent metals, Wigner and Seitz[21] took the potential in which a valence electron moves to be the potential $v_i(\mathbf{r})$ due to the ion in the polyhedron in which the electron is instantaneously located. The one-electron Schrödinger equation

$$[\mathbf{p}^2/2m+v_i(\mathbf{r})]\psi_\alpha(\mathbf{r})=\mathcal{E}_\alpha\psi_\alpha(\mathbf{r}) \tag{5.1}$$

was solved in the polyhedron subject to the Bloch boundary conditions at the surface of the polyhedron.[22] The sum of the energy parameters $\mathcal{E}_\alpha$ so obtained was found to be a very good approximation to the total energy of the metal, relative to the state in which the ions and electrons are infinitely dispersed.[23]

We now consider the effect of deforming the metal by a strain $\mathbf{s}$. The Schrödinger equation, (5.1), is unchanged, but the boundary conditions do change. We will repeat the device of changing coordinates, Eq. (3.5), so that the wave function expressed in the new coordinates satisfies the original boundary conditions. Expanding to first order in $\mathbf{s}$, we find

$$\delta\mathcal{E}_\alpha(\mathbf{s})=\int_{[0]}\psi_\alpha^*[-\mathbf{p}\cdot\mathbf{s}\cdot\mathbf{p}/m+\mathbf{r}\cdot\mathbf{s}\cdot\nabla v_i(\mathbf{r})]\psi_\alpha. \tag{5.2}$$

Transforming to a surface integral as before, we find

$$\delta\mathcal{E}_\alpha(\mathbf{s})=-(i\hbar/2m)\int_{[0]}\mathbf{dS}\cdot[(\nabla\psi_\alpha)^*\mathbf{r}\cdot\mathbf{s}\cdot\mathbf{p}\psi_\alpha-\psi_\alpha^*\nabla(\mathbf{r}\cdot\mathbf{s}\cdot\mathbf{p})\psi_\alpha]. \tag{5.3}$$

Expression (5.3) is seen to be the same as D_α, Eq. (4.10), except that (5.3) is calculated with the Wigner-Seitz wave functions and (4.10) is calculated with the exact wave functions.[24]

The result (5.3) is very interesting, considering that, in the simple Wigner-Seitz approximation, the sum of the $\mathcal{E}_\alpha$ is equal to the total energy of the metal. Unfortunately, the situation in a divalent metal is much more complicated so that no simple analogous statement can be made.

6. COMPARISON WITH THE WORK OF JONES AND GOODENOUGH

In this section we discuss the relationship of our work to that of Jones and Goodenough. We find, that with suitable assumptions, term D, Eq. (4.10), can be made to yield the estimates of Jones and Goodenough for the "effective stress."

Both authors base their estimates on the weak-potential model,[25] and neglect any changes in energy gaps with strain. Such a neglect in the weak-potential model is equivalent to assuming that the magnitudes of corresponding Fourier coefficients of the potential are unchanged by deformation. However, if the Fourier coefficients were actually constant, the potential would be "deformable;" that is, the potential after deformation is given in terms of the original potential by

$$v_s[(\mathbf{1}+\mathbf{s})\cdot\mathbf{r}]=v_0(\mathbf{r}), \tag{6.1}$$

where v_0 and v_s are the potentials before and after deformation.

Using Eq. (6.1) for the transformation of the potential and the method used before of changing variables to calculate the change in the one-electron energy eigenvalue, we obtain

$$\delta\mathcal{E}_\alpha=D_\alpha'=-\int_\Omega d^3r\psi_\alpha^*(\mathbf{p}\cdot\mathbf{s}\cdot\mathbf{p}/m)\psi_\alpha. \tag{6.2}$$

We have called this expression D_α' as it is analogous to D_α. It is to be noted that there is no potential energy term in (6.2). If the wave functions are expressed as linear combinations of plane waves, D_α' becomes

$$D_\alpha'=-\sum\mathbf{K}|a_\alpha(\mathbf{k}+\mathbf{K})|^2(\hbar^2/m)(\mathbf{k}+\mathbf{K})\cdot\mathbf{s}\cdot(\mathbf{k}+\mathbf{K}), \tag{6.3}$$

[16] J. Bardeen, Phys. Rev. **52**, 688 (1937).

[17] T. Koopmans, Physica **1**, 104 (1934). See also the discussion in F. Seitz, *Modern Theory of Solids* (McGraw-Hill Book Company, Inc., New York, 1940), p. 313.

[18] See, for example, R. W. Keyes, Phys. Rev. **92**, 580 (1953).

[19] E. N. Adams, Phys. Rev. **96**, 803 (1954).

[20] J. Bardeen and W. Shockley, Phys. Rev. **80**, 72 (1950). See also W. P. Dumbke, Bull. Am. Phys. Soc. **29**, No. 7, 15 (1954).

[21] E. Wigner and F. Seitz, Phys. Rev. **43**, 804 (1933); **46**, 509 (1934).

[22] Actually, Wigner and Seitz simplified the boundary conditions by replacing the polyhedron by a sphere. J. C. Slater [Phys. Rev. **45**, 794 (1934)] made a calculation in which he satisfied the Bloch boundary conditions at a few points on the surface of the polyhedron. In this section we shall assume that the exact boundary conditions have been satisfied.

[23] Wigner and Seitz also made a more elaborate calculation than the one described above. However, we will confine our attention to the simpler case.

[24] The fact that expression (5.3) gives the change in the Wigner-Seitz energy when the polyhedron is replaced by a sphere, was pointed out by Fuchs and Peng (reference 13). They used the expression to derive the formulas of H. Frohlich [Proc. Roy. Soc. (London) **A158**, 97 (1937)] and J. Bardeen [J. Chem. Phys. **6**, 372 (1938)] for the dependence on volume of the Wigner-Seitz energy of the state of wave number zero.

[25] L. Brillouin, *Quantenstatistik* (Verlag Julius Springer, Berlin, 1931).

where $a_\alpha(\mathbf{k}+\mathbf{K})$ is the coefficient of the normalized plane wave of propogation vector $\mathbf{k}+\mathbf{K}$, and the $\mathbf{K}$ are reciprocal lattice vectors multiplied by 2π.

In order to obtain the Jones and Goodenough estimates we assume that the wave functions are the nearly-free electron wave functions,[25] i.e., all a's are zero except a few corresponding to the smallest magnitudes of $\mathbf{k}+\mathbf{K}$. When $\mathbf{s}$ is given by (3.2), we distinguish three cases.

(1) If the end of the vector $\mathbf{k}$ lies at certain points of high symmetry, all the factors $(\mathbf{k}+\mathbf{K})\cdot\mathbf{s}\cdot(\mathbf{k}+\mathbf{K})$ have the same value. As the sum of the a's is unity (due to the normalization condition), we find

$$D_\alpha' = -(\hbar^2/m)\mathbf{k}\cdot\mathbf{s}\cdot\mathbf{k}. \tag{6.4}$$

The expression (6.4) is the change with strain of a free electron state with propogation vector $\mathbf{k}$, and thus is equal to the Jones estimate.

(2) For other points on the zone faces, the present method is different from that of Jones, though the two methods yield the same result in the weak potential case. To find the average change in energy of holes near points such as the corners of the zone V and W, Jones takes the change of free electron energy at the corners. [The change in free electron energy at point V is equal to the change in free electron energy at point W for the deformation (3.2).] However, the quantity D_α', evaluated for the state at W (or V), gives the change in energy of the state originally at W (or V). Under the deformation (3.2) the two states move away from points W and V, and, in fact, the values of D_α' associated with the two states are not equal. In order to find the average change in energy of holes originally near points V and W we must average the two values of D_α', weighting each value with the appropriate density of states. If we use the free electron densities of states for points V and W, we obtain the same estimate as Jones.

(3) In the case that the end of the vector $\mathbf{k}$ is near to but does not touch a zone face, Eq. (6.3) gives, besides the change in free electron energy, a contribution which depends on the variation of the a's with $\mathbf{k}$. The value of D_α' in this case works out to be the estimate of Goodenough, plus a term $\mathbf{k}\cdot\mathbf{s}\cdot\nabla_\mathbf{k}\mathcal{E}_\alpha$ which is zero when summed over the Fermi distribution as the tensor defined by (3.2) has no trace.

We are now able to state the assumptions which will lead to the theories of Jones and Goodenough. First of all, the effects considered by these authors are contained in term D. Further, they take only that part of D which would result if the self-consistent potential were "deformable." Finally, they evaluate the "deformable" part of D using nearly-free electron wave functions.

7. NUMERICAL APPLICATION TO BERYLLIUM

In this section we make numerical estimates of the various contributions to the "effective electron stress" in the case of pure beryllium. We chose beryllium as it is the only hexagonal close-packed metal for which one-electron wave functions have been calculated for states other than the central point in the Brillouin zone. The wave functions available are approximate Hartree wave functions calculated by Herring and Hill,[26] using the orthogonalized plane wave method.[27]§

TABLE II. Results of calculation of term D_α, the "kinetic stress."

α (State)	D_α (OPW) / D_α (free)	D_α (free)/e (ev)	[a] $\lvert a_F\rvert^2$
Γ_4^-	0.714	−22.4	1.03
Γ_3^+	0.984	−22.4	0.985
K_5	0.865	13.2	0.91

[a] $\lvert a_F\rvert^2$ is the absolute square of the coefficient of the nearly-free electron wave function in the orthogonalized plane wave function.

First, we discuss term D, Eq. (4.10). The individual contributions from each state, D_α, are fairly easily calculated with the orthogonalized plane waves. Each valence-electron wave function is expressed as a linear combination of plane waves and inner-shell wave functions, the coefficients chosen to minimize the valence-electron energy while keeping the valence-electron wave function orthogonal to the inner-shell wave functions. The inner-shell wave functions in the case of beryllium are very small (about 10^{-3} of the value at the atom) in the region between atoms where the surface of integration is located; therefore, we neglect the contributions of the inner-shell parts of the wave function to the "effective stress." Thus, we need only evaluate D_α with a combination of a few plane waves.

From the computational view, there are two reasons why the actual value of D_α will be different from the value predicted on the nearly-free electron model. We may think of the wave function as being made up of an inner shell part, a combination of plane waves which would be the wave function in the nearly-free electron approximation, and higher plane waves. The value of D_α is then different from that for free electrons both because of the contributions of the higher plane waves, and because the coefficient of the nearly-free electron wave function is different from unity.

We have evaluated D_α for several states in the Brillouin zone, the results being listed in Table II. The labeling of states is the same as in the Brillouin zone of Fig. 2. For simplicity, we express our results by giving the ratio of our value of D_α to the value obtained on the nearly-free electron model for the same point in the Brillouin zone. For each of the states investigated, the amplitude of the nearly-free electron part of the wave function is nearly unity (see Table II) and the amplitudes of the higher plane waves are small. As a result, the computed values of D_α resemble the free electron

[26] C. Herring and A. G. Hill, Phys. Rev. **58**, 132 (1940).
[27] C. Herring, Phys. Rev. **57**, 1169 (1940).
§ *Note added in proof.*—It has been brought to the author's attention that calculations have been performed for several points in the Brillouin zone for magnesium. See M. Trlifaj, Czech. J. Phys. **1**, 110 (1952).

values rather more than might have been expected *a priori*.

It should be noted that though the contribution to D from each of the states listed in Table II is similar in magnitude to the free electron contribution, the sum over the Fermi distribution could be very much different from the sum calculated using the free electron approximation. This is true because the total sum is a small difference between large terms, so that a deviation from the free electron value which is different in different parts of the zone could cause a large error in the sum. Such an error would, of course, cause a large error in the calculated value of the axial ratio.

From the information given in Table II, we can get some idea about the effect of strain on the energy gap. As Γ_4^- is the state just above the top face of the zone, and Γ_3^+ is the state just below, the difference between $D(\Gamma_4^-)$ and $D(\Gamma_3^+)$ would be the change in energy gap if one could neglect the effects of terms S and X. Table II shows that this change is of the order of thirty percent of the change in free electron energy.

We now consider the contribution of term S to the "effective stress." In order to make an estimate of S, we approximate the total charge density as that given by positive charges of $2e$ on the atomic positions and a uniform negative charge of $-2e$ per polyhedron. It is then convenient to transform S to a somewhat different form. It can be shown that

$$S(\omega)=\delta W(\omega)+(2\omega/\Delta)(4e/\Delta)\int_{[0]} d^3r\mathbf{r}\cdot\boldsymbol{\epsilon}\cdot\nabla\phi(\mathbf{r}), \quad (7.1)$$

where δW is the change with shear of the electrostatic energy of the simple model assumed above, and ϕ is the electrostatic potential of such a charge density.

The first term in Eq. (7.1), δW, can be calculated using the method of Ewald,[28] as developed by Fuchs.[29] That the variation of W with c/a is a factor in determining the axial ratio was pointed out by Herring and Hill.[26] We find

$$\delta W=-(2\omega/\Delta)0.00155e^2\epsilon/a. \quad (7.2)$$

It can be shown that if $\phi(\mathbf{r})$ were spherically symmetrical, the second term in (7.1) would be zero. We imagine expanding ϕ as a sum of multipoles. The largest contribution should come from the quadrupole term as the operator $\mathbf{r}\cdot\boldsymbol{\epsilon}\cdot\nabla$ has quadrupole symmetry when operating on a spherically symmetric function. The coefficient of the quadrupole term in the expression for ϕ has been calculated by Campbell, Keller, and Koenigsberg.[30] Replacing the polyhedron by a sphere of equal volume, we find for the second term in (7.1)

$$(4e/\Delta)\int_{[0]} d^3r\mathbf{r}\cdot\boldsymbol{\epsilon}\cdot\nabla\phi\cong 0.0025e^2\epsilon/a. \quad (7.3)$$

We then find for the approximate value of S in the case of beryllium 0.006ϵ electron volts per atom. Referring to Table I, we see that this value is small compared to the observed values of the "effective stress." Though our estimate of S is admittedly approximate, the small value obtained is an indication that the contribution of term S to the "effective stress" for a pure metal can be ignored.

Term X is very complicated and we have been able to make only a rough estimate of its value. If, in deriving the expression for the "effective stress," it is specified that the one-electron wave functions are solutions of the Hartree equations (instead of the Hartree-Fock equations), it is found that term X takes the somewhat different form,

$$X=-(e^2/2)\sum_{\alpha\beta}\int_\Omega d^3r_1\int_\Omega d^3r_2\psi_\alpha^*(\mathbf{r}_1)\psi_\beta(\mathbf{r}_1)$$
$$\times\psi_\beta^*(\mathbf{r}_2)\psi_\alpha(\mathbf{r}_2)(\mathbf{r}_1-\mathbf{r}_2)\cdot\boldsymbol{\epsilon}\cdot\nabla(1/|\mathbf{r}_1-\mathbf{r}_2|). \quad (7.4)$$

As the Herring and Hill wave functions are Hartree wave functions, expression (7.4) should be used in calculating X.

In evaluating (7.4), we neglect interaction between the valence electrons and closed shells[31] and replace the valence electron wave functions by plane waves. In such an approximation, the value of X depends sensitively upon the occupation of states in reciprocal space, being zero for a spherical Fermi surface. We attempt to represent the effect of the full zone which contains two electrons per atom by assuming the electrons occupy an ellipsoid of revolution which has the same volume and quadrupole moment [integrated value of $\frac{1}{2}(3k_z^2-k^2)$] as the zone. With the above assumption, we obtain an estimate of 0.1ϵ ev per atom for X for beryllium. Though the estimate is crude, it indicates that the exchange effect is important.

We are now in a position to try to make an estimate of the total "effective electron stress" in beryllium. Following Jones, we think of the total stress as that due to the full zone plus that due to overlap electrons minus that due to holes. We will approximate the term D for the full zone by summing the free electron value of D_α over the zone. Such a sum yields

$$\sum_{\alpha,s} D_\alpha\cong 0.8\epsilon \text{ ev per atom},$$

[28] See, for example, M. Born and M. Goeppert-Mayer, *Handbuch der Physik* (Verlag Julius Springer, Berlin, 1933), Vol. 24, Part 2, p. 710.

[29] K. Fuchs, Proc. Roy. Soc. (London) **A151**, 585 (1935).

[30] Campbell, Keller, and Koenigsberg, Phys. Rev. **84**, 1256 (1951).

[31] In beryllium the exchange interaction between closed shells on different atoms is negligible, but can be important in other metals. [See K. Fuchs, reference 29; also, Proc. Roy. Soc. (London) **A153**, 622 (1936); **A157**, 444 (1936).] However, to the extent that such interactions are represented by central forces between nearest neighbors only, they will produce no "effective stress" tending to change the axial ratio.

where the subscript z means that the sum is over the complete zone. Adding this to the estimate of X for the full zone made above, we have for the "effective stress" due to the full zone

$$\sigma_z \cong -0.9 \text{ ev per atom.}$$

The observed "effective stress" in beryllium is minus 0.44 ev per atom (see Table I), about one-half our estimate. If there occurred overlap at point K and holes at points V and W as the results of Herring and Hill suggest, the correction for the difference from a filled band would increase the calculated "effective stress" even more.

The above estimate of the "effective stress" is in error due to the partial use of the free electron approximation in obtaining it. Even so, the approximation is sufficiently good to yield an estimate for the "effective stress" in beryllium of the right sign and order of magnitude. For metals of larger atomic number whose wave functions are more complicated, it is to be expected that the free electron approximation is less accurate. In zinc and cadmium, the "effective stress" actually has the opposite sign to that predicted with the free electron approximation.

8. EFFECTS OF ALLOYING

In this section we discuss the change in lattice parameters caused by alloying. We find that the method used by Jones to calculate the change in axial ratio with alloying is incorrect, though we show that the correct method may lead to conclusions similar to those of Jones.

Besides changing the axial ratio, alloying causes changes in the volume, and produces local distortions of the crystal structure. We shall describe each configuration by giving the average axial ratio and average volume. For a given configuration the equilibrium positions of the nuclei are determined by the condition that the energy is a minimum subject to the constraint on the average axial ratio and average volume. We use the parameter ϵ as defined before, and introduce a parameter δ which is equal to one-third of the fractional difference between the average volume of the configuration and the volume of the pure metal at absolute zero.

The equilibrium values of ϵ and δ are those which minimize the total energy. The expansion (to second order) of the total energy in terms of ϵ and δ is

$$E = E_0 - \sigma\epsilon - \pi\delta + \tfrac{1}{2}\mu\epsilon^2 + \tfrac{1}{2}\nu\delta^2 + \eta\epsilon\delta. \tag{8.1}$$

Minimizing (8.1) with respect to ϵ yields

$$\epsilon_{\min} = (\sigma - \eta\delta)/\mu. \tag{8.2}$$

We could, of course, minimize the energy with respect to δ, and make a theoretical calculation of both ϵ and δ. However, we shall instead take δ from experiment, approximate μ and η by the appropriate combinations of elastic constants, and confine ourselves to the calculation of σ.

It is seen that the axial ratio changes with alloying because of three effects: the change in elastic constants, the change in atomic volume, and the change in the "effective stress."

We cannot say how large the effect due to the change of elastic constants is, since measurements of the relevant elastic constants for alloys of hexagonal metals have not been made. However, measurements on other systems indicate that the percentage change in elastic constants per atomic percent of solute varies from 0.5 to 5.[32] From Table I we see that the percentage change in ϵ_{obs} per atomic percent of solute is of the order of 10, so that the change in elastic constants may well have an important effect.

The change in axial ratio due to the change in atomic volume, given by the last term in (8.6), is the same as that derived by Jones using a somewhat different argument.

The change in the "effective stress" with alloying is due to the changes in the "effective phonon stress" and the "effective electron stress." We shall treat here only the change in the "effective electron stress" and neglect the change in the "effective phonon stress."

The "effective stress" is, as before, the negative of the derivative of the energy with respect to ϵ. A change in ϵ not only changes the average axial ratio but also affects the local distortions. We could express the actual deformation as a homogeneous deformation which changes the axial ratio, plus local deformations. The local deformations do not cause a first-order change in energy as the energy of the ideal configuration is a minimum with respect to local deformations. Thus we may obtain σ by calculating the linear change in energy caused by a homogeneous deformation.

We shall calculate the change in "effective electron stress" with alloying using the formulation discussed in Sec. 3, so that the total "effective stress" comes from the contributions of the surface atoms to B [Eq. (3.8c)].[33] We limit the discussion to the case of infinite dilution, so that there is a negligible number of impurity atoms near the surface. At first thought it might seem that the impurity atoms could not contribute to the surface stress due to the fact that the forces between atoms are short-ranged. However, the *stress field* due to the presence of the impurity atoms is long-ranged. In fact, the field at large distances from the impurity can be treated by classical elasticity theory, and falls off as the inverse cube of the distance from the impurity atom.

In calculating the excess in term B due to the presence of the impurities, we may replace the summation over μ by an integration over the surface of the sample, as the excess surface stress is a slowly varying

[32] R. A. Artman and D. O. Thompson, J. Appl. Phys. **22**, 358 (1951); A. D. N. Smith, J. Inst. Metals **80**, 477 (1952).

[33] The quantity $\langle \nabla_\mu V \rangle$ is zero for nuclei in the interior of the sample. If this were not so, the energy could be lowered by displacing interior atoms, but the energy of the ideal configuration is a minimum with respect to such displacements.

function. We find

$$\delta E=\delta E_0+\int_S (\mathbf{dS}\cdot\boldsymbol{\tau})\cdot\mathbf{s}\cdot\mathbf{r}, \tag{8.3}$$

where δE_0 is the linear change in energy for the pure metal and $\boldsymbol{\tau}$ is the stress tensor due to the presence of impurities. As the leading term in the stress varies as the inverse cube of the distance from the impurity, the second term in (8.3) will be proportional to the total number of impurities in the sample.[34] Thus the change in lattice parameters will be proportional to the concentration of impurities (for small concentrations). The stress field itself is determined by the atomic displacements near the impurity atoms. A complete theoretical calculation of the change in lattice parameters with alloying then involves a calculation of the local displacements.[35]

The results of the above discussion support the simple view of a metal as being made up of atoms which interact through short-range forces. That classical elasticity theory can be applied to the problem of change of lattice parameters has been recognized for a long time.[36] The importance of the present results is that they show that the simple "Brillouin zone effects" on the change of lattice parameters with alloying as hitherto calculated, do not exist. Note that there would be no change in average lattice parameters if there were not local distortions of the lattice, a fact in direct contradiction with the theory of Jones.

The reason for the failure of the Jones model can be elucidated by the following discussion. Let us calculate by our method the change with alloying of the contribution of term D [Eq. (4.2a)] to the "effective stress," remembering that such a contribution is analogous to the stress calculated by Jones. We shall further assume that there are no local distortions of the lattice. We choose the surface of ω to be far from impurity atoms. We then find that D is unchanged by the addition of the impurities. The reason is roughly this: the perturbation of the charge distribution due to an impurity atom dies off exponentially with distance from the impurity, so that the charge density at the surface of ω is the same as for the pure metal.[37] Thus, though there may be a contribution to the stress from the extra electrons introduced by alloying, it is exactly cancelled by the change in the stress due to the electrons already present.

In order to explain the experimental regularities in the change in lattice parameters it is now necessary to explain why there are regularities in the local displacements caused by alloying.[38] We present here a rough argument which indicates that such a program is possible. We assume that the behavior of the electrons near the impurity can be described in the effective mass approximation.[39] For the sake of argument, let us assume that the energy band structure is that proposed by Jones and discussed in Sec. 1 of this paper. There will be several types of states with energy near the energy of the Fermi level, corresponding to the different regions K, V, and Γ. Let us consider the region Γ. In analogy to the weak potential case, we may expect that the effective mass parallel to the c axis is smaller than that perpendicular to the c axis. Thus the orbit of an electron of effective mass corresponding to region Γ and which is bound to the impurity would extend further away from the impurity in the direction of the c axis than perpendicular to the c axis. Thus the contribution to the excess charge density from such electrons would extend farther up the c axis than perpendicular to it, causing a greater repulsive force on nuclei which lie in the direction of the c axis away from the impurity atom. Thus, electrons in states such as Γ tend to increase the axial ratio, the same conclusion as that drawn by Jones. Naturally, a sudden onset of overlap at Γ would cause a change in the excess charge density, which in turn would effect the lattice parameters. The same reasoning gives a conclusion qualitatively the same as Jones's concerning the region K, though the effect of the holes at points V will depend on details of the band structure which are not as easily ascertained.

9. SUMMARY AND CONCLUSIONS

The main purpose of this paper was to investigate the bases of the theories of Jones and Goodenough concerning the change in axial ratios of hexagonal crystals with alloying. As explained in the introduction, the theories are interesting as they seem to afford easy ways of getting information about the band structure. In order to elucidate these theories we carried out two separate investigations which, surprisingly enough, turned out to be rather unrelated.

The first investigation (Secs. 3 through 7) consisted of examining the factors responsible for the deviation of the axial ratios of pure metals from the ideal close-packing value. The major factor found was the "kinetic stress," which resembles the stresses estimated by Jones and Goodenough. In fact, in beryllium the contributions of each one-electron state to the "kinetic stress" are

[34] Actually, there is also a constant term in the stress tensor due to impurities. In the method of this section the magnitude of the constant term is determined by the condition upon the average axial ratio and average volume. A more straightforward procedure would be to apply originally the boundary condition that there are no surface stresses and determine the change in lattice parameters from the surface displacements.

[35] For a discussion of the calculation of local atomic displacements in several cases, see D. L. Dexter, Phys. Rev. **87**, 768 (1952).

[36] B. Ya. Pines, J. Phys. (U.S.S.R.) **3**, 309 (1940).

[37] Actually, it is necessary that the contribution to the charge density at the surface of ω from each group of states with the same energy is unchanged by alloying. See J. Friedel, Advances in Physics **3**, 446 (1954).

[38] Another kind of argument to show that the "effective sizes" of solute atoms (and thus local displacements) depend upon the relative valency of the solvent and solute has been made by J. H. O. Varley, Phil. Mag. **45**, 887 (1954).

[39] For a discussion of the validity of the effective-mass approximation as applied to impurity states, as well as a discussion of the character of the exact solutions, see G. F. Koster and J. C. Slater, Phys. Rev. **95**, 1167 (1954).

numerically very much like those given by the Jones theory. It is pointed out, however, that the sum of the contributions from all states could be very much different. Two other factors tending to change the axial ratio were found, one of which, the exchange stress, could be appreciable. A rough estimate made of the "effective stress" in beryllium was of the right order of magnitude.

The second investigation concerned the change of the lattice parameters with alloying. It was found that the change in axial ratio with alloying is due to a different mechanism from that proposed by Jones. However, use of the effective-mass theory allowed us to establish a connection between the band structure and the change in axial ratio when the correct mechanism is taken into account. It is remarkable that using the correct method, we draw the same qualitative conclusions concerning band structure as does Jones! It may be necessary to use a more sophisticated method in order to obtain numerical agreement with experiment, but the effective-mass treatment doubtlessly contains the essential physics. Thus, the most attractive feature of the Jones theory remains, i.e., the possibility of getting information about the band structure from the alloying data.

In the course of the first investigation a formalism was developed which is capable of giving the first-order effect of a homogeneous deformation on the energy band structure. Several possible applications of such a formalism were mentioned in Sec. 4.

The author wishes to express his sincere gratitude to Dr. E. N. Adams II for his invaluable advice and encouragement. Thanks are also due Dr. Conyers Herring for making available copies of the wave functions.

PHYSICAL REVIEW VOLUME 98, NUMBER 2 APRIL 15, 1955

Properties of Germanium Doped with Nickel

W. W. Tyler, R. Newman, and H. H. Woodbury
General Electric Research Laboratory, Schenectady, New York

(Received December 22, 1954)

The temperature dependence of electrical resistivity in p- and n-type nickel-doped germanium crystals indicates that nickel introduces two acceptor levels in germanium at 0.22 ± 0.01 ev from the valence band and 0.30 ± 0.02 ev from the conduction band. Ionization energies deduced from infrared photoconductivity studies at 77°K are in agreement with the values obtained from resistivity measurements. N-type samples show higher photosensitivity than p-type samples and demonstrate quenching effects. The distribution coefficient for nickel in germanium is about 2.3×10^{-6}.

I. INTRODUCTION

THIS paper presents measurements of electrical and optical properties of germanium crystals containing nickel as the major impurity. Results are similar in most respects to those reported previously for Fe-doped[1,2] and Co-doped[3] germanium crystals. In a paper primarily concerned with the effects of Ni and Cu on carrier lifetime in germanium, Burton *et al.*[4] have reported low-temperature conductivity and Hall coefficient data which indicated that Ni introduces an acceptor level at about 0.23 ev from the valence band of germanium. This conclusion is confirmed. In addition, evidence is presented indicating that Ni introduces another acceptor level at about 0.30 ev from the conduction band. A summary is given comparing the properties of Fe-, Co-, and Ni-doped germanium crystals.

[1] W. W. Tyler and H. H. Woodbury, Phys. Rev. **96**, 874 (1954).
[2] R. Newman and W. W. Tyler, Phys. Rev. **96**, 882 (1954).
[3] Tyler, Newman, and Woodbury, Phys. Rev. **97**, 669 (1955).
[4] Burton, Hull, Morin, and Severiens, J. Phys. Chem. **57**, 853 (1953).

II. EXPERIMENTAL RESULTS

A. Crystal Preparation

The methods of crystal growth and sample preparation used with Ni-doping experiments were similar to those used in previous work with Fe- and Co-doping.[1,3] Experimental techniques used in studying the Ni-doped samples have also been described.[1,2]

Johnson Matthey Company (JM) "spectroscopically pure" Ni and Sigmund Cohn (SC) "chemically pure" Ni were used for doping. Initial experiments using the JM Ni indicated that about 10 percent of the carriers introduced into the Ge crystals as a result of doping were probably due to low ionization energy acceptor impurities in the Ni. The SC Ni proved somewhat more satisfactory either because of a lower concentration of specific low ionization energy acceptor impurities or because acceptor impurities were compensated by donor impurities. Spectroscopic analyses of both sources of Ni indicated "barely detectable" amounts of B and "traces" of Al. The SC Ni analysis also indicated a

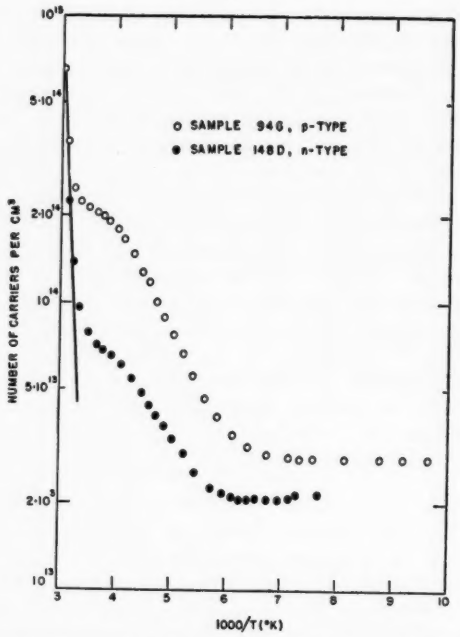


FIG. 1. Temperature dependence of carrier concentration for low-resistivity n- and p-type Ni-doped crystals. (On the ordinate, $2\cdot10^{3}$ should be $2\cdot10^{13}$.)

"slight trace" of Ga.[5] In an attempt to decrease the concentration of these acceptor impurities, single crystals of Ni were grown using the SC Ni as raw material. Comparative spectrographic analyses were made of the starting material and a third generation Ni crystal grown from central portions of preceeding crystals. There was no evidence of significant reduction of the B, Al, or Ga content from spectroscopic analyses or from doping experiments.

Subsequent to initial doping experiments which indicated the presence of low-ionization-energy acceptor impurities in the Ni, all crystals were grown with small amounts of dilute As-Ge or Sb-Ge alloys added to the melt to compensate acceptor impurities. Because of their greater homogeneity, the most satisfactory counter-doping alloys for this purpose are Ge single crystals grown with suitable As- or Sb-doping. In

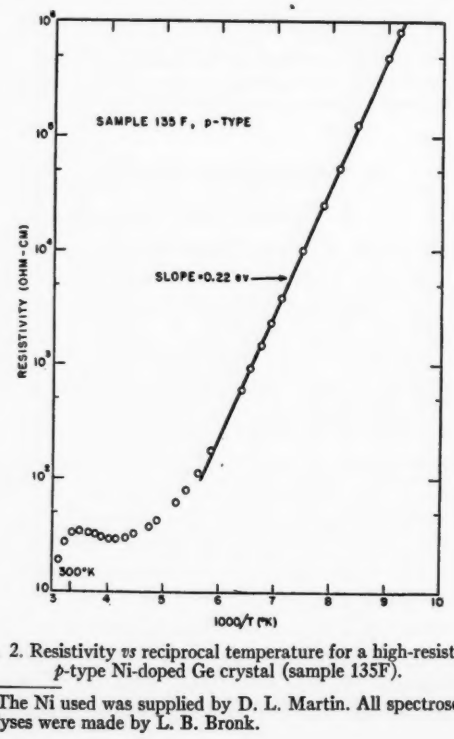


FIG. 2. Resistivity vs reciprocal temperature for a high-resistivity p-type Ni-doped Ge crystal (sample 135F).

[5] The Ni used was supplied by D. L. Martin. All spectroscopic analyses were made by L. B. Bronk.

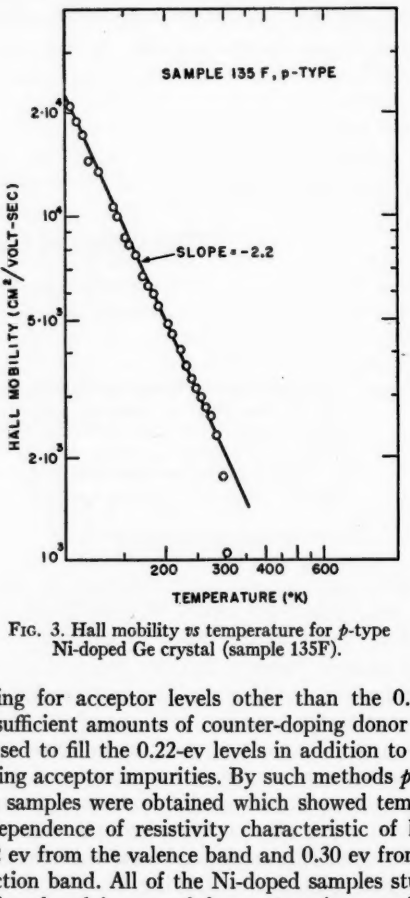


FIG. 3. Hall mobility vs temperature for p-type Ni-doped Ge crystal (sample 135F).

searching for acceptor levels other than the 0.22-ev level, sufficient amounts of counter-doping donor alloy were used to fill the 0.22-ev levels in addition to compensating acceptor impurities. By such methods p- and n-type samples were obtained which showed temperature dependence of resistivity characteristic of levels at 0.22 ev from the valence band and 0.30 ev from the conduction band. All of the Ni-doped samples studied could be placed in one of four categories, consistent with the two-acceptor model. Assume N Ni atoms per cm³, p holes per cm³ due to the acceptor impurities in the Ni, and n electrons per cm³ from the counter-doping alloy. If $n-p<0$, the sample is low-resistivity p-type. Due to freeze-out of the holes supplied by Ni atoms, the resistivity first rises on lowering the temperature but becomes limited by the low-ionization-energy ac-

ceptor states. If $0<n-p<N$, the sample is high-resistivity p-type and shows a temperature dependence of resistivity characteristic of the lower Ni acceptor state. If $N<n-p<2N$, the sample is high-resistivity n-type and shows a temperature dependence of resistivity characteristic of the upper Ni level. For $2N<n-p$, a low-resistivity n-type sample is observed with the resistivity rise limited by excess low-ionization-energy donors.

B. Electrical Properties

Figure 1 shows plots of number of carriers *vs* reciprocal temperature for Ni-doped Ge crystals characteristic of the first and last categories described above. The number of carriers was obtained from the Hall coefficient using the relation $R_H=1/ne$. From these curves, values of N can be obtained and estimates made of the effective distribution coefficient for Ni in Ge. Values obtained are 2.4×10^{-6} for sample 94G and 2.2×10^{-6} for sample 148D. These values are about a factor of 2 lower than the value given by Burton *et al.*[4] The principal source of error for the values given above is lack of exact knowledge of the concentration of Ni in the melt during crystallization of the sample. The effective distribution coefficient may also depend to some extent on the concentration of Ni in the melt, on rate of crystallization and rate of cooling after crystallization. For 94G, the concentration in the melt was 1.9×10^{-3} and for 148D about 5.7×10^{-4}. In each case the rate of crystallization was about 1.4 grams per minute. The agreement between values of effective distribution coefficient for these p- and n-type crystals indicates that the number of upper acceptor levels introduced by Ni is equal to the number of lower levels within experimental error.

Figures 2 and 3 show the temperature dependence of resistivity and mobility for a high-resistivity p-type Ni-doped sample. Eight such p-type samples from these different crystals were studied. Energy values given by the slopes of resistivity *vs* reciprocal temperature range from 0.21 ev to 0.23 ev. Figures 4 and 5 show similar curves for a high-resistivity n-type Ni-doped sample. Six such samples from six different crystals were studied. Energy values range from 0.28 ev to 0.31 ev.

The temperature dependence of Hall mobility shown in Figs. 3 and 5 is not typical of all the samples studied. In many high-resistivity Ni-doped samples, particularly for n-type samples, anomalous Hall mobility effects were observed. Figure 6 shows measured mobility data for high-resistivity n- and p-type samples 152A and 142I. Resistivity data for these samples showed the characteristic energies of 0.30 ev and 0.23 ev, respectively, throughout the temperature range studied. The anomalous mobility behavior generally becomes pronounced at temperatures below 200°K. Very few high-resistivity samples showed well-behaved mobility down to 100°K. In some cases the measured mobility values would break from the normal linear relationship

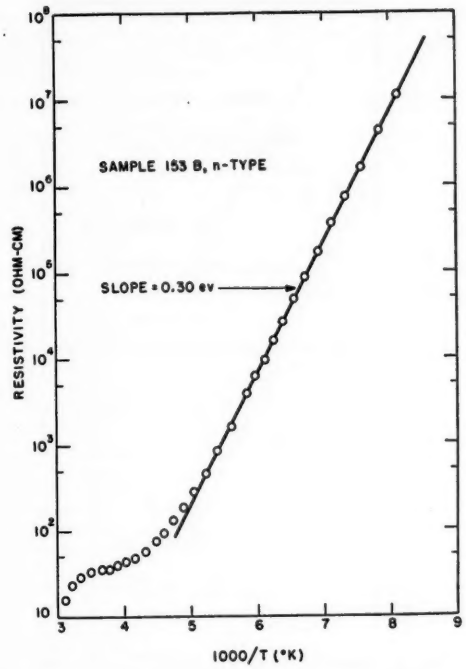


FIG. 4. Resistivity *vs* reciprocal temperature for a high-resistivity n-type Ni-doped Ge crystal (sample 153B).

and become relatively independent of temperature; in other cases mobility values would fall very rapidly after breaking from the normal relationship, sometimes

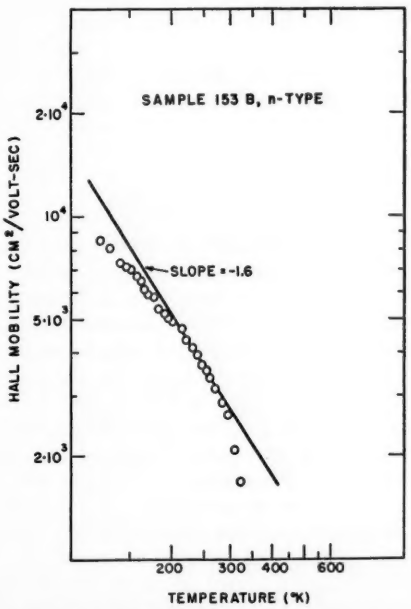


FIG. 5. Hall mobility *vs* temperature for n-type Ni-doped Ge crystal (sample 153B).

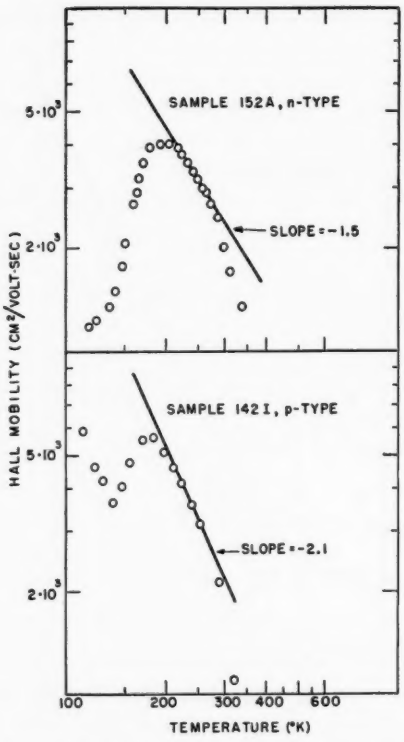


FIG. 6. Anomalous Hall mobility of high-resistivity n- and p-type Ni-doped Ge crystals (samples 152A and 142I).

falling several decades in a temperature interval of 30 to 40°K. Anomalous mobility behavior observed in high-resistivity Ni-doped samples is thought to be due to macroscopic potential variations which develop as the samples are cooled. Resistivity may not rise uniformly throughout the sample and small disconnected regions in the sample may actually remain at a relatively low resistivity level.

C. Photoconductivity

Infrared photoconductivity studies at 77°K were made on p-type samples from three different crystals and on n-type samples from three different crystals. Figures 7 and 8 show the spectral dependence of photoconductivity for typical p- and n-type samples. The curves are qualitatively similar to those reported previously for Au[6]-, Fe[2]-, and Co[3]-doped Ge crystals. In the region of impurity photoconduction the photosensitivities of n- and p-type Ni-doped samples were comparable. In the region of intrinsic absorption ($h\nu > 0.7$ ev), n-type samples were a factor of 10 to 10^2 more photosensitive than p-type samples. Quenching effects in n-type samples were observed. Maximum quenching effects occurred at about 0.5 ev with observed decreases in background photoconduction of about 50 percent. The spectral dependence of photoconduction in either p- or n-type samples was very reproducible from sample to sample and seemed independent of mobility behavior. Anomalous mobility behavior presumably affects the magnitude but not the spectral dependence of photoconduction in samples which are all of one type. For one Ni-doped sample, a spectral response was observed which behaved like n-type material from 0.70 to 0.45 ev and showed the characteristic p-type spectrum from 0.45 to 0.22 ev. Careful probing studies in the dark at low temperature subsequently confirmed that this sample was of mixed type.

Room-temperature lifetime measurements were made on many of the Ni-doped samples. Lifetime values were obtained from the rate of photoconductivity decay after spark excitation. Samples grown from a melt containing about 0.08 atomic percent Ni with hole concentrations of about 10^{14} per cm³ indicated carrier lifetime of several microseconds or less. The results are qualitatively similar to those of Burton *et al.*[4]

III. SUMMARY AND DISCUSSION

Table I presents a summary of impurity levels in Ge crystals which have been identified with the presence of Fe, Co, and Ni impurities in the melt. Approximate values for effective distribution coefficients have also been listed. E_{vi} represents energy measured from the valence band to the impurity level in p-type samples; E_{ci} is measured from the impurity level to the conduction band in n-type samples. The experimental error

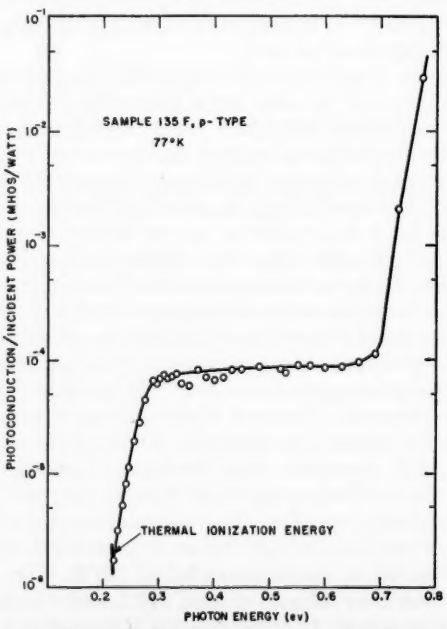


FIG. 7. Photoconductive spectrum of a p-type Ni-doped Ge crystal at 77°K (sample 135F).

[6] R. Newman, Phys. Rev. **94**, 278 (1954).

in determining slopes of resistivity plotted against reciprocal temperature, from which the energy values of impurity levels were obtained, is ± 0.02 ev at most. The use of resistivity data rather than Hall coefficient data in determining energy values is somewhat arbitrary. According to the simple theory that leads to a 3/2 power dependence of mobility on temperature, the dependence of mobility on temperature cancels the theoretical change in effective density of states with temperature, justifying the use of resistivity data in determining energy values. The mobility does not follow the 3/2 power law in p-type material but correction of energy values to account for this would raise the quoted E_{vi} values by only about 0.01 ev. Such corrections have not been carried out because experimental and theoretical uncertainties in interpretation of the data are comparable in magnitude. No information has been obtained showing possible temperature dependence of impurity levels.

Photoconductivity measurements made at 77°K are qualitatively similar for high-resistivity Fe-, Co-, and Ni-doped Ge crystals. Ionization energies deduced from impurity photoconduction thresholds in n- and p-type samples are consistent with the values quoted in Table I, except for n-type Fe-doped samples, for which photoconductivity studies indicate a somewhat deeper level, similar to n-type Co- and Ni-doped crystals. In general, n-type samples are more photosensitive than p-type samples, exhibit quenching effects and are slower in response to changes in light levels.

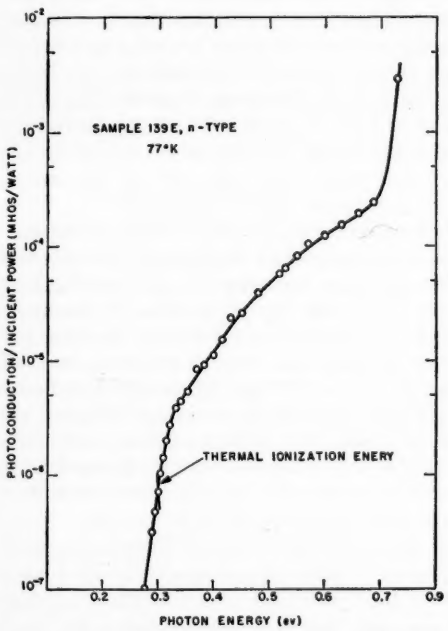


FIG. 8. Photoconductive spectrum of an n-type Ni-doped Ge crystal at 77°K (sample 139E).

TABLE I. Impurity levels and distribution coefficients (k_0) for iron, cobalt, and nickel in germanium. E_{vi} represents energy measured from the valence band to the impurity level in p-type samples; E_{ci} is measured from the impurity level to the conduction band in n-type samples.

	E_{vi} (ev)	E_{ci} (ev)	k_0
Fe	0.34	0.27	$<10^{-6}$
Co	0.25	0.31	$\sim 10^{-6}$
Ni	0.22	0.30	$\sim 2.3\times 10^{-6}$

It is tentatively assumed that Fe, Co, and Ni are substitutional impurities in the Ge lattice, each acting as a double acceptor giving rise to the energy levels quoted in Table I. Evidence that the impurity states are acceptors is direct for Ni and Co and is assumed for Fe because of similar atomic properties of Fe, Co, and Ni and the analogous photoconductive properties of crystals doped with Fe, Co, and Ni. No evidence that these impurities introduce more than two deep levels has been obtained.

Evidence for the double-acceptor model is not considered conclusive. It is possible that the two different levels observed for each of the transition metals studied are due to impurity atoms at two different types of sites at about the same concentration. The number of lattice imperfections in even the best crystals is probably sufficient to account for the observed solubility of Fe, Co, and Ni. Observed similarity of the impurity photoconductive spectra of n-type crystals may imply that the upper levels are due to structural defects or some other property common to Fe-, Co-, and Ni-doped crystals and independent of the atomic properties of the specific doping impurity.

Most of the uncertainty concerning the properties of Ge crystals doped with transition metals is related to their low effective distribution coefficients and the presence, in the purest metals available, of significant concentrations of high distribution coefficient impurities. Also, in samples dominated by high-ionization-energy impurities the carrier concentration at low temperatures decreases to very low values and it is difficult consistently to obtain homogeneous samples with well-behaved mobility. This problem has been particularly serious for Ni-doped samples. Uneven cooling of the ingot after crystallization, combined with a pronounced retrograde solid solubility could be in part responsible for this lack of homogeneity.

Stopping Cross Sections of Metals for Protons of Energies from 400 to 1000 kev*

D. Wayne Green,† John N. Cooper, and James C. Harris
The Ohio State University, Columbus, Ohio

(Received December 13, 1954)

The stopping cross sections of manganese, copper, germanium, selenium, silver, tin, antimony, gold, lead, and bismuth are reported for protons in the energy range 400 to 1000 kev. The cross sections are roughly proportional to the square root of the atomic number of the stopping element and inversely proportional to the velocity of the incident protons. There is some evidence that the stopping cross section increases more rapidly than $Z^{\frac{1}{2}}$ as the various p shells are filled.

I. INTRODUCTION

STOPPING power has been the subject of interest and extensive investigation since the earliest days of nuclear physics.[1] From almost the very beginning the range of an energetic charged particle has been used as a measure of its energy. These range-energy relations were first experimentally investigated by observing alpha particles from natural radioactive sources. With the development of the particle accelerators it became possible to obtain protons and other charged particles with well defined energy. The development of techniques for obtaining thin foils permitted the direct observations of the rate of energy loss as a function of energy rather than its inference from range measurements.

Theoretical aspects of the problem have been investigated by a number of physicists. Of particular importance are the early treatments by Bohr[2] using a classical approach and by Bethe[3] and Bloch[4] using the quantum mechanical approach. The expressions obtained have been reasonably successful in predicting the experimental results for higher energies. However, one of the basic assumptions involved in these theories is that the velocity of the particle is greater than the orbital velocity of the atomic electrons of the stopping material. As a consequence, the results for low particle velocities have not been entirely satisfactory. Several attempts have been made to correct this deficiency,[5,6] but none has proved completely adequate. Therefore, experimental data must be relied upon in this velocity range.

Of more immediate concern here are the expressions derived by Bohr[7] and more recently by Lindhard and Scharff[8] using a statistical approach based on the Thomas-Fermi model of the atom. While these forms are not expected to give accurate results, they do predict a relationship between the stopping power, the atomic number, and the velocity of the particles. The testing of these relations for various energies and stopping media is desirable. In this paper, data on the stopping cross sections of the elements manganese, copper, germanium, selenium, silver, tin, antimony, gold, lead, and bismuth are reported and compared with these theoretical predictions.

II. EXPERIMENTAL TECHNIQUE

The experiments were carried out using protons accelerated by a Van de Graaff generator. The acceleration voltage was stabilized by detecting an unbalance in proton current intercepted by defining slits and automatically altering the corona drain to bring the beam back to a balanced position. The energy of the protons was determined from current in the coils of the analyzing magnet which was calibrated by the use of the lithium and fluorine (p,γ) resonances.[9]

The method used was a modification of the one suggested and developed by Madsen and Venkateswarlu.[10] In this technique, an excitation spectrum for a proton reaction with some target material is obtained first without, and then with a foil of a stopping medium inserted in the proton beam. The energy loss suffered by the protons in the stopping material is indicated by the apparent shift in the resonance peaks of the reaction spectrum. This technique suffers a serious limitation in the fact that most materials do not lend themselves to the formation of foils which are sufficiently thin, and yet strong enough to be suitable.

To overcome this difficulty a technique was developed[11] in which the stopping material was evaporated directly onto the target material. (This modification was proposed and used independently by Madsen[12]

* Supported in part by the U. S. Atomic Energy Commission through a contract with The Ohio State University Research Foundation.

† Now at Knox College, Galesburg, Illinois.

[1] S. K. Allison and S. D. Warshaw, Revs. Modern Phys. **25**, 779 (1953), and references.

[2] N. Bohr, Phil. Mag. **25**, 10 (1913).

[3] H. A. Bethe, Ann. Physik **5**, 325 (1930).

[4] F. Bloch, Ann. Physik **16**, 285 (1933).

[5] M. S. Livingston and H. A. Bethe, Revs. Modern Phys. **9**, 245 (1937).

[6] J. O. Hirschfelder and J. L. Magee, Phys. Rev. **73**, 207 (1948).

[7] N. Bohr, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd. **18**, No. 8 (1948).

[8] J. Lindhard and M. Scharff, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd. **27**, No. 15 (1953).

[9] Hornyak, Lauritsen, Morrison, and Fowler, Revs. Modern Phys. **28**, 356 (1950).

[10] C. B. Madsen and P. Venkateswarlu, Phys. Rev. **74**, 648 and 1782 (1948).

[11] Green, Cooper, and Harris, Phys. Rev. **94**, 764 (1954).

[12] C. B. Madsen, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd. **27**, No. 13 (1953).

although he confined its application to those materials which would also produce foils.) This arrangement quite naturally gives rise to the name "sandwich target" since the target material is sandwiched between the target backing and the stopping material. For this work, LiF was used exclusively as the target material and tantalum was generally used for the target backing. The resonance energy levels for the (p,γ) reactions are accurately known[9] for both lithium and fluorine.

One of the complications introduced by the use of sandwich targets is the problem of determining accurately the thickness of the target material (LiF). The displacement of a resonance level is greater than that due to the stopping material by an amount equal to one half the energy thickness of this LiF layer. Since the target thicknesses used were in the neighborhood of five kev, it became necessary to take this into account. By using a calibration target evaporated at the same time as the sandwich target, the effect of the target thickness cancels out on the assumption that the LiF is the same thickness on both targets.

The necessity of making a simultaneous calibration run each time a sandwich target was run required a switching apparatus in the vacuum system which permitted inserting either a calibration target or a sandwich target in the proton beam as desired. The target assembly formed part of the Faraday cage which was connected to an integrator to measure the charge incident on the target. The gamma rays from the target were detected by a cylindrical 1.5×1.5 inch NaI crystal and a 5819 photomultiplier tube.

III. TARGETS

The general method of preparing targets was to evaporate lithium fluoride on the target blank, weigh the blank, evaporate the stopping material onto the lithium fluoride, and weigh the target again. The blanks were circular with a diameter of 0.5 inch. Tantalum was the usual material, although a few were copper or silver. The weighings were done on a microbalance which gave weights consistently reproducible to better than ten micrograms. With this sensitivity, it was necessary to deposit at least one milligram of stopping material to keep the weighing error under one percent.

TABLE I. Summary of results on stopping power for protons.

Proton energy (kev)	Stopping power (kev/mg/cm²) Mn	Cu	Ge	Se	Ag	Sn	Sb	Au	Pb	Bi
400	218	179	177	168	150	142	146	96	100	104
450	206	172	168	160	142	135	138	92	95	99
500	196	165	159	154	135	128	132	88	90	94
550	186	158	153	148	128	122	125	84	86	90
600	177	152	147	143	121	116	119	80	82	86
650	169	147	142	138	116	111	113	78	79	82
700	162	142	138	134	111	107	108	75	75	79
750	156	137	133	130	107	103	104	73	73	75
800	151	133	130	126	103	100	101	71	70	73
850	147	128	126	122	101	97	98	69	68	70
900	143	125	123	120	98	94	96	67	67	69
950	140	122	119	117	96	91	93	65	65	67
1000	136	119	116	115	94	89	91	64	64	66

TABLE II. The atomic stopping cross section (ev cm²×10⁻¹⁵) of elements for protons as a function of proton energy.

Proton energy (kev)	Element Mn	Cu	Ge	Se	Ag	Sn	Sb	Au	Pb	Bi
400	19.9	18.9	21.3	22.0	26.9	28.0	29.5	31.4	34.4	36.1
450	18.8	18.2	20.2	21.0	25.4	26.6	27.9	30.1	32.7	34.4
500	17.9	17.4	19.2	20.2	24.2	25.2	26.7	28.8	31.0	32.7
550	17.0	16.7	18.4	19.4	22.9	24.0	25.2	27.5	29.6	31.3
600	16.1	16.0	17.7	18.7	21.7	22.8	24.1	26.2	28.2	29.9
650	15.4	15.5	17.1	18.1	20.8	21.9	22.8	25.5	27.2	28.5
700	14.8	15.0	16.6	17.6	19.9	21.1	21.8	24.5	25.8	27.4
750	14.2	14.5	16.0	17.0	19.2	20.3	21.0	23.9	25.1	26.1
800	13.8	14.0	15.7	16.5	18.5	19.7	20.4	23.3	24.1	25.4
850	13.4	13.5	15.2	16.0	18.1	19.1	19.8	22.6	23.4	24.3
900	13.1	13.2	14.8	15.7	17.6	18.5	19.4	21.9	23.0	24.0
950	12.8	12.9	14.3	15.3	17.2	17.9	18.8	21.3	22.4	23.3
1000	12.4	12.6	14.0	15.1	16.8	17.5	18.4	21.0	22.0	23.0

The evaporations were carried out in a vacuum of 10^{-4} mm of mercury or better. The evaporating baskets were made of 0.020-in. tungsten wire wound in a conical shape with a diameter of 0.5 inch at the open end. The wire was coated with a paste of levigated alumina which was allowed to dry and then baked in a vacuum by passing a current through the coil. During evaporation the targets were mounted about six inches above the basket. The lithium fluoride was deposited on five target blanks at a time. The central blank of this cluster was used as the lithium fluoride calibration target for the other four which were subsequently made into "sandwich" targets. The stopping material was evaporated onto two of the blanks at a time. Comparison of the average thickness of the stopping material of the two targets gave an indication of the uniformity of the deposit. In most instances the two values agreed within two percent.

There are several reasons for believing that the thickness of the stopping material on a given target blank was sufficiently uniform so that thickness variations did not make a major error contribution. These thickness variations quite naturally fall into two categories: microscopic fluctuations covering an area smaller than that spanned by the proton beam and macroscopic ones extending over an area larger than that covered by the beam. Small fluctuations of the first type do not introduce serious errors in stopping cross section since the average energy loss corresponds closely to the average thickness. However, microscopic variations will produce a very noticeable effect in the straggling. Most of the data were not suitable for detailed straggling analysis, but rough calculations indicate that the non-uniformity never exceeded five percent and was less in a majority of cases.

Macroscopic fluctuations, on the other hand, can introduce serious errors. Macroscopic uniformity was carefully checked by running a few of the targets several times, making sure that the beam struck a different area of the target each time. The results of this in-

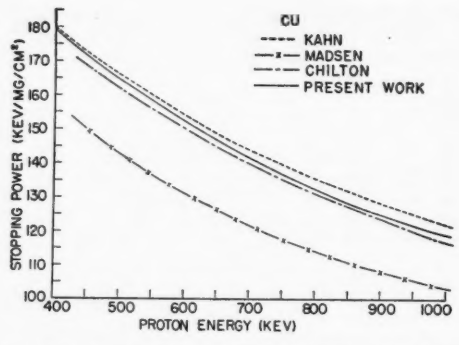


FIG. 1. The stopping power of copper for protons as a function of proton energy. Also plotted are the curves reported by Madsen (see reference 12), by Kahn (see reference 14), and by Chilton, Cooper, and Harris (see reference 13).

vestigation were consistent with the information obtained from the variation in average thickness previously referred to; namely, that the maximum thickness change from one side of the target to the other was about two percent. Since the targets were ordinarily positioned so that the beam struck the center of the target, the actual thickness should correspond closely to the average thickness. Further evidence of target uniformity comes from the internal consistency of the data from one stopping layer to another (for both foils and sandwich targets). Agreement was within five percent, except in the case of selenium.

In the study of copper and gold, foils were also used. They were made by evaporating the metal onto a microscope cover glass which was weighed before and after the evaporation. The foil was floated off the glass with distilled water and mounted on a frame which was then used in the same apparatus as has been described by Chilton, Cooper, and Harris.[13] Results from the foils agreed very well with those obtained from the sandwich targets.

The silver, tin, lead, and part of the gold used were obtained from the Fischer Scientific Company. The other materials were obtained from the Chemistry Department of The Ohio State University. All were of cp quality. As a further check, a semiquantitative spectrographic analysis was run on one target of each material after it had been bombarded. No significant amount of impurities was found in any of them. The analysis would not be expected to show carbon contamination.

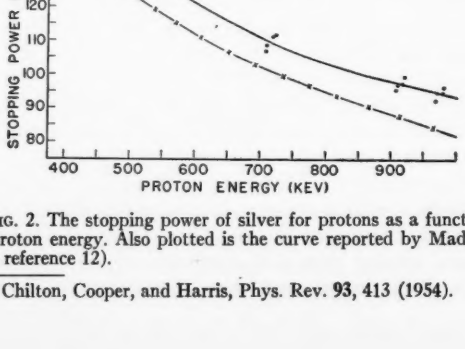


FIG. 2. The stopping power of silver for protons as a function of proton energy. Also plotted is the curve reported by Madsen (see reference 12).

[13] Chilton, Cooper, and Harris, Phys. Rev. **93**, 413 (1954).

IV. RESULTS

At least four separate targets were used for each of the stopping materials studied. The bombardment of each target was limited to about one microampere-hour to minimize the effect of carbon formation resulting from bombardment with the beam. Foils however, were run approximately twice this long as a rule. Each target run usually gave the stopping power at five different proton energies so the curve for each material was determined with a minimum of 18 to 20 points. Curves were drawn to make the best fit of the points, and the values appearing in Table I were read from these curves. The atomic stopping cross sections were computed from the data of Table I and are listed in Table II.

TABLE III. Target thicknesses in mg/cm².

Au	Sb	Sn	Mn	Cu	Ag	Pb	Bi	Se	Ge
1.16[a]	0.79	1.00	0.82	0.74	0.98	1.07	1.10	0.79	0.72
1.22[a]	1.07	0.78	0.80	0.72	0.74	1.04	1.40	1.34	0.70
0.84	1.08	0.99	0.75	0.86	0.93	1.74	1.10	1.28	0.81
0.85	0.79	0.83	0.78	0.89	0.77	1.78	1.43	0.81	0.81
1.23[a]				0.98[a]			0.88	0.88	
							0.84	0.92-0.75	

[a] Indicates foil.

The stopping power dE/dx determined from the displacement of a resonance level, was assigned to the average energy of protons in the stopping material. Chilton[13] has derived a correction to this average energy but in all cases the correction was so small that it could be neglected without increasing the probable error significantly.

The probable error for the curves is estimated to be about 2.5 percent, except in the case of selenium where it is about four percent. The energy loss suffered by the protons in the stopping material ranged from 70 kev to 150 kev in most cases and the peak positions could be read to better than one kev, so the error involved in the energy loss determinations was approximately one percent. The area of the targets was known to 0.25 percent. The weight of the stopping material ranged from 0.9 mg to 1.7 mg (see Table III) and the weight was accurate to about 0.01 mg which again gives an error of approximately one percent. The effect of other factors, such as carbon deposition during bombardment and non-uniformity of stopping layers, was checked several times but little if any contribution was noted except possibly in the case of selenium. The 2.5 percent

error estimate is supported by the fact that the maximum spread in the points for any curve was five to six percent.

VI. DISCUSSION AND CONCLUSIONS

The stopping power of copper was studied primarily as a check on the method used for this work, since it already has been reported by Kahn,[14] Chilton,[13] and Madsen.[12] The curves of Kahn and Chilton are about four to six percent apart while Madsen's curve is about 15 percent below them. The present curve lies between the curves of Kahn and Chilton as shown in Fig. 1.

Silver has also been studied by Madsen in the same energy region as the present work. As shown in Fig. 2, our curve lies about eight to ten percent above his.

In the case of gold, most of the data were obtained from three foils. Two sandwich targets were also used. Data from the two types of stopping layers were consistent. The curve shown in Fig. 3 lies about ten percent above Kahn's curve at one Mev and about 25 percent above it at 400 kev. Madsen's results are in essential agreement with Kahn. This is a very disturbing discrepancy since it is far greater than the experimental errors involved. No explanation for the disagreement is advanced.

TABLE IV. Product of stopping power and a factor proportional to proton velocity.

Proton energy (kev)	Element Mn	Cu	Ge	Se	Ag	Sn	Sb	Au	Pb	Bi
400	138	114	112	106	94.8	90.0	92.1	60.8	63.2	65.6
500	137	115	113	109	95.2	90.5	92.9	61.9	63.6	66.6
600	137	118	114	111	93.8	90.2	91.9	62.3	63.7	66.4
700	136	119	115	112	93.0	89.2	90.8	62.7	63.1	65.7
800	135	119	116	112	92.2	89.3	90.2	63.0	62.8	65.0
900	137	118	116	113	93.0	89.2	90.6	63.4	63.1	65.0
1000	136	119	116	115	93.5	88.8	91.2	63.9	64.2	66.2

According to the theories of both Bohr[7] and Lindhard and Scharff,[8] the atomic stopping power σ should be proportional to the reciprocal of the proton velocity. If this is correct, then the product of the velocity and σ should be a constant for a given element. This product is shown in Table IV for the ten elements on which measurements are reported. The product is seen to be roughly constant (within experimental error) for all elements except selenium. In the case of selenium, there was evidence of some physical change, possibly a carbon contamination, which gradually increased the effective thickness of the stopping material.

The theory of Bohr and the theory of Lindhard and Scharff are in disagreement on the dependence of stopping cross section on the atomic number. According to Bohr's theory, σ should be proportional to the cube root of the atomic number for higher values of Z while the theory of Lindhard and Scharff predicts that it should be proportional to the square root. To investigate this, a plot was made of the logarithm of σ as a function of the logarithm of Z. (See Fig. 4.) Included in this plot are data from Kahn,[14] from Dunbar, Reynolds, Wenzel, and Whaling,[15] from Chilton,[13] and from the present work. In cases of duplication where no serious discrepancy occurred, values obtained in this laboratory were used. For krypton, xenon, and gold, two values were included. The slope of the drawn line from hydrogen to argon is about 0.7, while the slope of the line making the best fit from argon to bismuth is about 0.5 as predicted by Lindhard and Scharff.

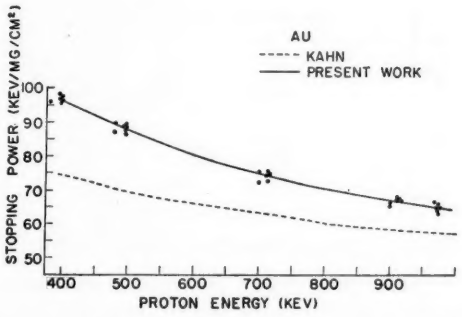


FIG. 3. The stopping power of gold for protons as a function of proton energy. Also plotted is the curve reported by Kahn (see reference 14).

In Fig. 5, a similar plot is shown of the region from argon to bismuth to give greater detail. Here there seems to be strong evidence of a periodic fluctuation of σ. From hydrogen to argon electrons are added to outer s and p shells, but with scandium the filling of the $3d$ shell beings. Between copper and krypton, electrons are again being added to fill the outer $4s$ and $4p$ shells. The light line on Fig. 5 from copper to krypton has a

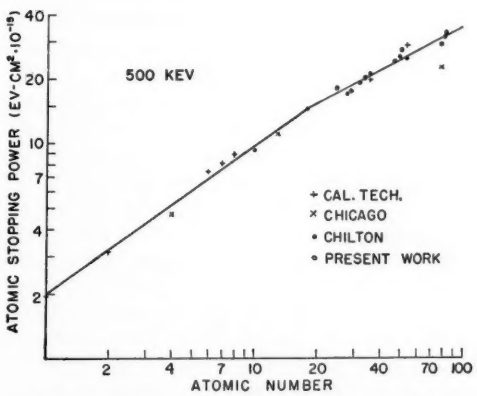


FIG. 4. The atomic stopping cross section σ as a function of atomic number for protons of 500-kev energy. In addition to the data of the present paper there are plotted points from the data reported by Kahn (see reference 14) of the University of Chicago, by Dunbar, Reynolds, Wenzel, and Whaling (see reference 12) of the California Institute of Technology, and by Chilton, Cooper, and Harris (see reference 13).

[14] D. Kahn, Phys. Rev. **90**, 503 (1953).

[15] Dunbar, Reynolds, Wenzel, and Whaling, Phys. Rev. **92**, 742 (1953).

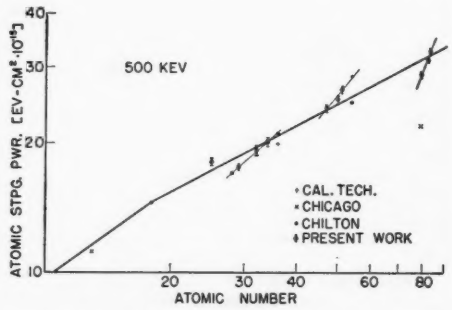


FIG. 5. The atomic stopping cross section σ as a function of atomic number, showing the relatively pronounced rise as the s and p shells are filled.

slope not much different from the line from hydrogen to argon. The same pattern is repeated starting with silver as the $5s$ and $5p$ levels are filled and again starting with gold as the $6s$ and $6p$ electrons are added. However, before too much credence can be given to this periodicity, measurements will have to be made on many more elements. It is interesting that Sternglass[16] has observed a similar periodic variation in the yield of secondary electrons as a function of atomic number when elements are bombarded with high-energy electrons. This is another phenomenon which involves the interaction between the atomic electrons and a high speed charged particle.

It should be noted that if the periodic slope changes as shown in Fig. 5 prove to be valid, then the dependence of the stopping power on the conductivity properties or upon the physical state (solid or gaseous) of an element is probably relatively small.

The authors wish to thank the other members of the Van de Graaff generator program for their help and to express their appreciation to Dr. A. B. Chilton and Professor Ward Whaling for valuable discussions. Financial support from The Ohio State University Development Fund and from the University Committee on Research Grants is gratefully acknowledged.

[16] E. J. Sternglass (private communication), and Phys. Rev. **80**, 925 (1950).

PHYSICAL REVIEW VOLUME 98, NUMBER 2 APRIL 15, 1955

Cosine Interaction in CsF and RbF*

G. BEMSKI,† W. A. NIERENBERG, AND H. B. SILSBEE
University of California, Berkeley, California

(Received October 7, 1954)

A series of measurements of the line shape of the fluorine resonance *versus* magnetic field have been made using a molecular-beam apparatus. Good theoretical agreement has been found for RbF. The CsF behavior is complicated by the overlapping Cs spectrum. The equivalent field at the fluorine nucleus per unit molecular rotation, H_r, is found to be 2.28 ± 0.20 gauss and 3.00 ± 0.15 gauss for RbF and CsF respectively. These results are less than those previously measured.

I. INTRODUCTION

EARLIER investigations of the fluorine resonance in the alkali fluorides resulted in discrepancies in the determination of the rotational constant c/h which is a measure of the strength of the spin-orbit coupling of the nuclear spin to the molecular rotation. The discrepancies were particularly clear for CsF and LiF. The molecular-beam magnetic-resonance apparatus (MBMR)[1] yielded 18 kc/sec for Li^6F^{19} compared to the value 37 kc/sec for the electrical apparatus[2] (MBER). Similarly, experiments[3] with $Cs^{133}F^{19}$ showed values of 14 ± 2 kc/sec (strong field, MBER) or 14 ± 3 kc/sec (weak field, MBER) compared to 8 kc/sec[1] (MBMR). In addition, the curve shape of the zero field MBMR resonance for CsF did not agree with that predicted from the strong field results. It was therefore decided to repeat the CsF experiments using MBMR including measurements at intermediate fields and to perform a similar experiment using RbF with the possibility of comparison to MBER.[4]

The earlier work was carried out at either high or low fields. Since that time a theory[5] has been conconstructed which predicts the shape of the resonance for all values of the magnetic field in terms of just one parameter. It was hoped that some clue to the discrepancies might be found in the transition from the extreme Zeeman to the complete Paschen-Back conditions. In addition, the verification of the theory has its own intrinsic interest.

II. EXPERIMENTAL DETAILS

Since this experiment was the first to be completed using the new molecular beam apparatus at Berkeley, a brief description of the apparatus will be attempted.

* This research has been supported by the Office of Naval Research.

† Now at the Bell Laboratories, Murray Hill, N. J.

[1] W. A. Nierenberg and N. F. Ramsey, Phys. Rev. **72**, 1075 (1947).

[2] J. C. Swartz and J. W. Trischka, Phys. Rev. **86**, 606 (1952).

[3] J. W. Trischka, Phys. Rev. **74**, 718 (1948).

[4] V. Hughes and L. Grabner, Phys. Rev. **79**, 314 (1950).

[5] W. A. Nierenberg, Phys. Rev. **82**, 932 (1951).

The apparatus is similar to its predecessors at Columbia and Harvard and was operated in a conventional manner for these experiments.[6,7]

The vacuum envelope was constructed of $\frac{1}{4}$-in. stainless steel, 12 in. in diameter, and 80 in. long. This provided for improved rigidity and no variation in beam intensity was observed due to instability of the can. The three magnets were machined from Armco iron, the pole tips conforming to the equipotentials of a two wire system.[8] The field gradient was 10^5 gauss/cm at a field of 12 000 gauss. The geometry was such that the deflection for an average alkali fluoride molecule due to a change in orientation of the F^{19} nucleus was approximately 0.006 in. at the detector. The oscillating fields were 28 cm long and therefore had a limiting resolution of approximately 1500 cycles/sec for the alkali fluorides. However, since the lines studied were of the order of half a megacycle per second in width, this resolution was not needed.

The data described in this article were obtained as averages of the differences of four beam readings with the oscillator off and five with the oscillator on. No significant variation in the results was noticed over a range of rf current between 2.5 and 5 amperes.

The oven slit-width, the collimator slit-width, and the detector diameter were all 0.001 in. The surface ionization detector fed a conventional DuBridge-Brown electrometer circuit employing a 5800 Victoreen electrometer tube.

III. THEORY OF THE EXPERIMENT

The usual Hamiltonian employed in describing this experimental situation is

$$\mathcal{H}=-\mu_N g_I \mathbf{I}\cdot\mathbf{H}-\mu_N g_J \mathbf{J}\cdot\mathbf{H}-c\mathbf{I}\cdot\mathbf{J}. \qquad (1)$$

$\mathbf{I}$ is the fluorine spin in units of $\hbar$; $\mathbf{J}$ is the molecular angular momentum in units of $\hbar$; μ_N is the nuclear magneton; $\mathbf{H}$ is the applied magnetic field; c is the magnetic interaction in ergs per unit rotation of the molecule with the fluorine nucleus; g_I is the magnetic moment for fluorine in units of μ_N; g_J is the molecular magnetic moment in units of μ_N. In addition, there are similar terms for the other nucleus, a quadrupole term for the other nucleus and a dipole-dipole term between the two nuclei.

Both because $\bar{J}\gg 1$ and $g_J \ll g_I$ in these experiments, a very good approximation is to treat $\mathbf{J}$ as a stationary, classical vector. This implies[5] that in observing the fluorine resonance we may neglect the terms in the Hamiltonian involving the alkali nucleus. The dipole-dipole term is neglected because its contributions may be estimated as too small (~4 kc/sec) for the effects considered here. With these approximations, the predicted[5] line shape is:

$$\frac{dN}{dy}=\frac{\sqrt{\pi}}{2}\frac{y}{y_I}[\mathrm{erf}(y+y_I)-\mathrm{erf}|y-y_I|], \qquad (2)$$

where

$$y=(ha/c)\nu, \quad y_I=(ha/c)\nu_I, \quad a=(\sqrt{\pi})/2\bar{J},$$

$$\mathrm{erf}(x)=(2/\sqrt{\pi})\int_0^x \exp(-\xi^2)d\xi.$$

$\bar{J}$ is the average over the Boltzmann distribution for the molecules in the oven, relating a to the molecular moment of inertia K through $a=(h\sqrt{\pi})/(2\pi kKT)^{\frac{1}{2}}$. ν is the applied radiofrequency; y is therefore a dimensionless variable. ν_I is the Larmor frequency of the fluorine nucleus; y_I is therefore a dimensionless parameter. The quantity ha/c is the one parameter that determines the shape of the resonance curves for all values of the field.

The observed transitions are those due to $\Delta m_I=\pm 1$, $\Delta m_J=0$ in strong fields or $\Delta m_F=0,\pm 1$ in weak fields. Transitions $\Delta m_J=\pm 1$, $\Delta m_I=0$ are not detectable in this apparatus.

In strong fields the curve is symmetric and has a sharp peak centering on the Larmor frequency. As the field is decreased, the resonance curve becomes asymmetric and at vanishingly small fields the position of the maximum is determined by the internal interaction alone.

IV. THE EXPERIMENTAL RESULTS

Figure 1 and Tables I and II display the data. The fluorine resonance has been observed for fields ranging

TABLE I. Fluroine resonance in RbF.

H gauss	ν_I Mc/sec	Peak position Mc/sec	$\frac{1}{2}$ width Mc/sec	c/ha Mc/sec	y_I
0	0.00	0.4	0.68	0.60	0.00
70	0.28	0.4	0.55	0.50	0.57
125	0.50	0.6	0.63	0.61	0.82
170	0.68	0.68	0.58	0.58	1.17
220	0.88	0.90	0.56	0.56	1.57
328	1.31	1.35	0.63	0.63	2.08
500	2.0	2.0	0.60	0.60	3.33
750	3.0	3.0	0.55	0.55	5.46
975	3.9	3.9	0.62	0.62	6.30

Average value of $c/ha=0.58$ Mc/sec.

TABLE II. Fluorine resonance in CsF.

H gauss	ν_I Mc/sec	Peak position Mc/sec	$\frac{1}{2}$ width Mc/sec	c/ha Mc/sec	y_I
0	0.00	...	...	...	0.00
130	0.52	0.51	0.46	0.45	1.15
165	0.66	0.66	0.60	0.60	1.10
990	3.96	3.96	0.71	0.71	5.6

High-field value of $c/ha=0.71$ Mc/sec

[6] Rabi, Millman, Kusch, and Zacharias, Phys. Rev. **55**, 728 (1939).
[7] J. B. M. Kellogg and S. Millman. Revs. Modern Phys. **18**, 323 (1946).
[8] Rabi, Kellogg, and Zacharias, Phys. Rev. **96**, 157 (1934).

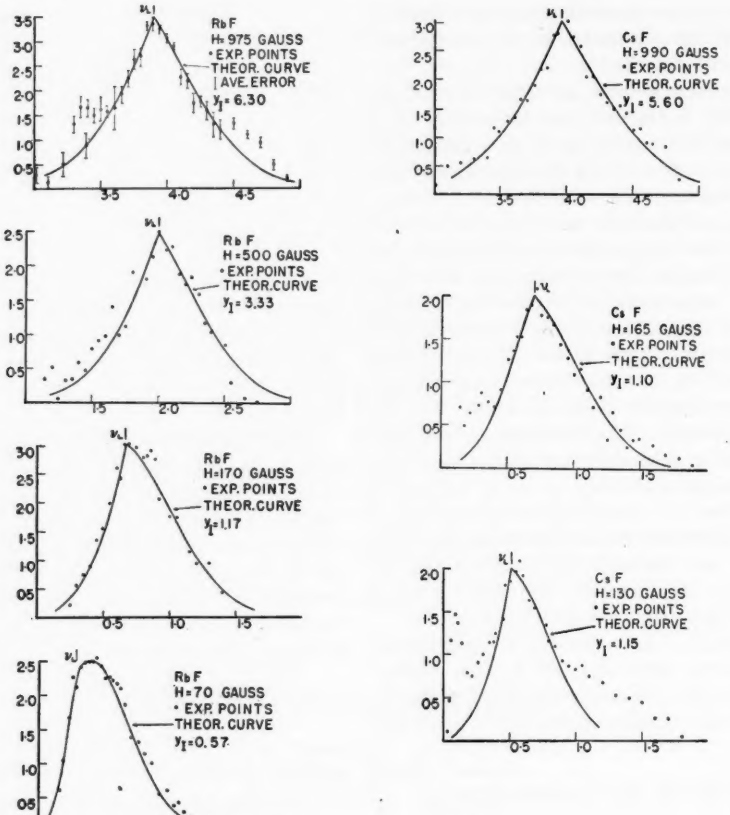


FIG. 1. The fluorine resonance (a) in RbF, and (b) in CsF. The ordinates are percent reduction in beam intensity; the abscissas are frequencies in megacycles per second. The solid lines are the theoretical curves; the solid points are the experimental results. The typical statistical error is shown on the RbF curve for $H=975$ gauss. Additional data for RbF at $H=750$, 328, and 125 gauss are not shown because of space limitations, but the curves are in no essential way different from the ones shown.

from strong ($H>500$ gauss, $y_I>1$), through intermediate ($300>H>100$ gauss, $y_I\sim1$) to weak ($H<100$ gauss, $y_I<1$). Figure 1(a) shows the resonance curves for RbF *vs* magnetic field. The points are the experimental results. The solid lines are the best theoretical fits. It should be emphasized that the entire family of experimental results for the complete range of field are fitted by a theory that has but one adjustable parameter and that parameter may be estimated independently. The agreement between theory and experiment is excellent and there would seem to be little question that the assumed Hamiltonian, th theory, and the approximations made apply very wel to RbF. The quantity ha/c is the only parameter fo these fields and frequencies which is needed to explai these results. Unfortunately the moment of inertia o RbF is not well known. Schomacker and Stevenson's[9] method for calculating the nuclear separation yield 2.54 A for RbF. This value of r in addition to the ove temperature value of 975°K for RbF fixes $\bar{J}=56.3$

[9] V. Schomacker and D. P. Stevenson, J. Am. Chem. Soc. 63, 3 (1941).

This corresponds to a most probable $J=45.9$ which agrees with the value of 46 used by Zeiger and Bolef.[10] The average value of $c/ha=0.58$ Mc/sec combined with $a=0.0157$ results in the value $c/h=9.1\pm0.9$ kc/sec. The principal uncertainty is in the value of the calculated r. The experimental uncertainty is less by a factor two. The related field per unit molecular rotation at the position of the fluorine nucleus $H_r=2.28\pm0.20$ gauss.

Figure 1(b) is the corresponding family of resonance curves for CsF. Here the agreement with experiment over the range of C field is poor. The line shape at very strong fields is in good agreement with theory. At intermediate and weak fields the disagreement is marked. In Sec V, most of the discrepancy is explained but as a result only the value $c/ha=0.71$ Mc/sec for strong fields can be used to evaluate c/h. The moment of inertia of CsF has been determined by Trischka[11] as $K=(152\pm6)\times10^{-40}$ g cm^2 and the corresponding nuclear separation as $r=2.36\times10^{-8}$ cm. (These values are for the first vibrational state. The values for other vibrational states differ too little to make an appreciable difference in this calculation.) As a result $J=52.4$, $a=0.0169$, and $c/h=12.0\pm0.6$ kc/sec. The principal source of error is in the present experiments. The field H_r is 3.00 ± 0.15 gauss.

The maximum drop in beam intensity varied between 2 and 3.5 percent. This variation depends on the half-width of the beam at the detector. This, in turn, depends on the effective width of the oven slit which had a tendency to broaden with time due to "oven plugging." The CsF beam in particular gave considerable trouble and the beam was often three times wider than predicted by the geometry. The very large g factor of the fluorine nucleus was the chief reason that a sufficiently large effect was obtained. The errors in determination of c/ha are due to:

1. The actual fitting of the theoretical curve of proper half-width. A possible error of 2–3 percent is not excluded.

2. Average error in measurement of beam changes in intensity. The 975-gauss curve, Fig. 1(a), shows this vertical spread in detail.

3. Changes in magnetic field values during a run. The error on this account was less than 1 percent for $H>100$ gauss, and 2 percent for $H<100$ gauss.

4. Frequency. The frequency was measured to better than 1 percent.

5. Absolute error in C field. A flip-coil calibration checked the sharp peak of the fluorine resonance at high fields to better than 2 percent. A residual field exists for $H<200$ gauss and the field values reported have been corrected for this effect. The corrections do not affect the c/ha values.

V. DISCUSSIONS AND CONCLUSIONS

It has been previously noted (Sec. IV) that the RbF spectral curves fit the theory with excellent agreement. The result, $c/h=9.1\pm0.9$ kc/sec is less than but consistent with Hughes and Grabner's[4] result of 12 ± 3 kc/sec. (This number is obtained by taking the weighted average of 11 ± 3 for $Rb^{85}F$ and 14 ± 3 for $Rb^{87}F$.) These measurements were made at Columbia on the electric molecular apparatus by using low rotational states. At first glance the agreement with theory is surprising inasmuch as the Rb resonance could be expected to confuse the pattern. However, this resonance has been extensively investigated by Zeiger and Bolef.[10] At zero magnetic field the maximum observable resonance due to the Rb^{87} is confined to the region 8.0–8.6 Mc/sec and has a peak intensity of 0.2 percent due to the large quadrupole moment. The F resonance at zero field is confined to the region 0.0–1.5 Mc/sec and has a peak intensity of 2.6 percent. Therefore the Rb^{87} spectrum is ineffectual at zero field for two reasons. At higher fields where the two spectra may overlap, the Rb^{87} intensity is reduced even further by the Zeeman spreading but the F resonance intensity is essentially unchanged. Zeiger and Bolef never observed the Rb^{85} resonance, presumably because the peak intensity was considerably less than 0.2 percent. This result has a sound theoretical basis.

The CsF strong field result $c/h=12.0\pm0.6$ kc/sec must be compared with Trischka's[11] results for strong and weak fields of 18 ± 2 and 14 ± 3 kc/sec respectively. The agreement with the average, 16 ± 3 kc/sec is poor and the deviation is in the same direction. The poor agreement of the line shape at intermediate and weak fields [Fig. 1(b)] has been tentatively assigned to the Cs resonance. The measured[11] $e^2qQ/h=0.62$ Mc/sec. Cs has a spin of 7/2. Under these circumstances Feld and Lamb[12] predict zero-field quadrupole peaks at 0.0444, 0.0888, and 0.1332 Mc/sec. These peaks are sufficient to explain the anomalous behavior of the low-frequency side of the F resonance. The departure from the high-frequency side of the F resonance is not understood.

[10] D. I. Bolef and H. J. Zeiger, Phys. Rev. **85**, 799 (1952).
[11] J. W. Trischka, Phys. Rev. **74**, 418 (1948).
[12] B. T. Feld and W. E. Lamb, Jr., Phys. Rev. **67**, 15 (1945).

Equilibrium Charge Distribution of Energetic Nitrogen Ions

H. L. REYNOLDS, L. D. WYLY,* AND A. ZUCKER
Oak Ridge National Laboratory, Oak Ridge, Tennessee
(Received November 1, 1954)

The equilibrium charge distribution of nitrogen ions in Formvar was measured as a function of the nitrogen energy from 6 to 26 Mev. The fringing magnetic field of the cyclotron was used to separate the charge groups resulting from the passage of the nitrogen ion beam through Formvar foils of sufficient thickness to insure an equilibrium charge distribution. Over the energy range investigated, the ionic charge is distributed primarily between four, five, six, and seven electron charges. The percentage of completely stripped ions varies from 42 percent at 26 Mev to 6 percent at 11 Mev. The ratio of the electron loss to capture cross section between N^{7+} and N^{6+} as well as between N^{6+} and N^{5+} is proportional to $V^{4.0}$, whereas between N^{5+} and N^{4+} the ratio is proportional to $V^{2.9}$. This ratio is unity for N^{6+} and N^{7+} at approximately 28.2 Mev, and for N^{5+} and N^{6+} at 13.3 Mev.

INTRODUCTION

As an energetic heavy ion passes through matter it undergoes a process of pickup and loss of electrons. A criterion commonly used is that the pickup and loss cross sections for a particular electron are nearly equal when the velocity of the ion is equal to the bound state velocity of that electron. This criterion has been fairly well established for alpha particles and protons. The fact that ions pick up electrons was first established by Henderson[1] when he observed energetic singly charged alpha particles from a radioactive source. A survey of electron exchange data for alpha particles and protons has been given by Allison and Warshaw.[2] Results for heavier ions have been summarized by Knipp and Teller,[3] but the data they use are limited to low energy particles which were obtained from alpha-particle recoils. Stier[4] and others have studied the charge exchange process for heavy ions with energies of less than 1 Mev. The theoretical aspects of charge exchange have been discussed by Bohr.[5] The problem of the charge of fission fragments in low-pressure gases has been treated by Bell.[6]

Reynolds, Scott, and Zucker[7] have measured the *average* charge of nitrogen ions in nickel as a function of energy. A better understanding of the charge exchange process could be gained if the equilibrium charge distribution (the fraction of the particles with a particular charge after equilibrium has been reached) as a function of energy could be obtained. This information would also be of use to experimenters who desire multiply-charged ion beams for accelerators. Recently, Stephens and Walker[8] have measured the equilibrium charge distribution for nitrogen ions at 15 Mev.

In the present investigation, the equilibrium charge distribution of nitrogen ions has been measured from 6 to 26 Mev by utilizing the fringing magnetic field of the cyclotron to separate the charge groups. The equilibrium charge was obtained in thin Formvar films.[9] A preliminary report of the measurement at 26 Mev has been published.[10]

EXPERIMENTAL METHOD

The triply charged nitrogen ions in the deflected beam of the 63-inch cyclotron were passed through a foil and separated, according to their charges, by the fringing magnetic field of the cyclotron. A separation of 2 mm between adjacent charge groups was obtained with a path length of approximately 8 in. (the average field is approximately 1500 gauss and is non-uniform). Therefore, it was necessary to collimate the beam highly and to minimize multiple scattering in the foils so that the charge groups at the detector would not overlap. This was a difficult criterion to fulfill because of the appreciable scattering of nitrogen ions due to their large charge.

The apparatus used is illustrated in Fig. 1. The distance from slit 1 to slit 2 is $3\frac{5}{16}$ in. and the distance from slit 2 to the detector is 8 in. The entire apparatus shown was open to the cyclotron vacuum and was located in the fringing magnetic field. The initial beam passed through slit 1, which is 0.001 in. wide by 0.5 in. long. This slit was constructed of carbon so that gamma rays produced by the nitrogen striking the carbon could be detected with a Geiger counter placed nearby. These gamma rays were used to monitor the beam. Nickel foils were placed directly across the first slit to reduce the energy of the beam to the desired value. After passing through the first foil the beam reached its equilibrium charge distribution and was widely divergent in direction due to multiple scattering. The second slit, 0.0015 by 0.5 in. and made of 0.0015-in. thick aluminum, was movable from outside the vacuum;

* Permanent address: School of Physics, Georgia Institute of Technology, Atlanta, Georgia.

[1] G. H. Henderson, Proc. Roy. Soc. (London) **A102**, 496 (1922).
[2] S. K. Allison and S. D. Warshaw, Revs. Modern Phys. **25**, 779 (1953).
[3] J. Knipp and E. Teller, Phys. Rev. **59**, 659 (1941).
[4] Stier, Barnett, and Evans, Phys. Rev. **96**, 973 (1954).
[5] N. Bohr, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd. **18**, 8 (1948).
[6] G. I. Bell, Phys. Rev. **90**, 548 (1953).
[7] Reynolds, Scott, and Zucker, Phys. Rev. **95**, 671 (1954).
[8] K. G. Stephens and D. Walker, Phil. Mag. **45**, 543 (1954).
[9] Formvar E resin was obtained from Shawenegan Products Corporation, 350 Fifth Avenue, New York.
[10] H. L. Reynolds and A. Zucker, Phys. Rev. **95**, 1353 (1954).

it was used to select and collimate a portion of the beam. A thin foil, usually Formvar, 50 $\mu g/cm^2$ thick,[11] was placed on the second slit. It was in this foil that the final equilibrium charge distribution was reached. A thin foil was used in this position to minimize the multiple scattering. The analyzed beam at the top of the chamber was detected with photographic emulsions. The photographic plate was in a container covered with a 0.2 mg/cm^2 aluminum foil to prevent light from reaching the emulsion during the loading and unloading of the plate.

Since a magnetic field existed between foils 1 and 2, the different charge groups produced in foil 1 entered foil 2 at different angles due to the different radii of their paths. Each charge group entering foil 2 was, again, broken up into the equilibrium charge distribution. If the initial beam was separated into an equilibrium condition of three different charges and each of these were again separated into three charge groups, the result at the detector would be nine lines, if they did not overlap. To prevent this confusing possibility, a magnetic shield was placed between foils 1 and 2 so that the particles entered foil 2 at the same angle, independently of their charge. The magnetic shield consisted of two concentric soft iron cylinders separated by a brass spacer. The outer cylinder had an outer diameter of 3 in. and a thickness of $\frac{7}{16}$ in. The inner cylinder had an outside diameter of $1\frac{5}{8}$ in., and a thickness of $\frac{3}{16}$ in. The magnetic field inside the shield was thus reduced to the extent that charges were not noticeably separated between slits 1 and 2.

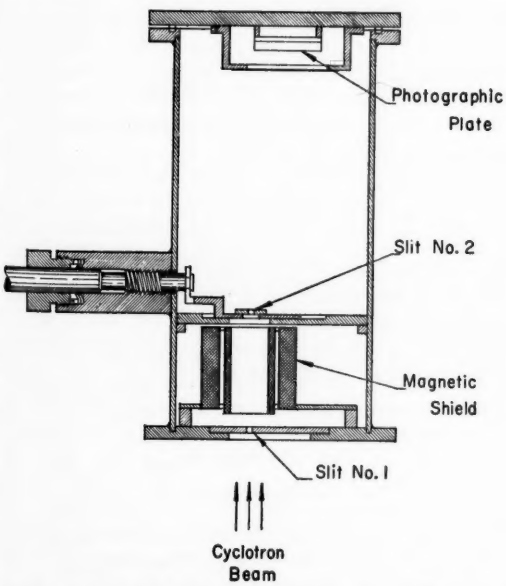


FIG. 1. Schematic drawing of chamber used to collimate and analyze nitrogen beam. The magnetic field is perpendicular to the paper.

[11] We wish to thank Dr. B. H. Ketelle for providing us with the thin Formvar foils used in this experiment.

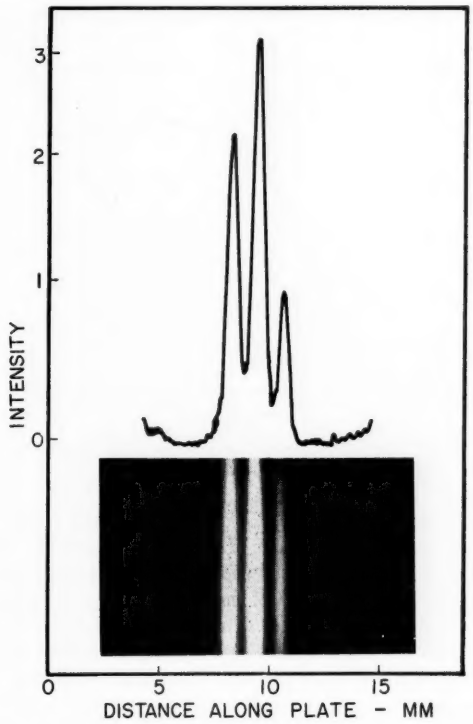


FIG. 2. An enlarged photograph of a typical exposure taken at 23.9 Mev. The photodensitometer trace is shown above the photograph. Beginning at the left the ionic charges are seven, six, and five.

Ilford C-2 emulsions, 50 microns thick, were used as detectors; exposures were usually less than one second. The second slit was aligned with the aid of a zinc sulfide screen in place of the plate holder. The density of the lines was read with a Leeds and Northrup traveling-stage micro-photodensitometer. An example of the line structure obtained in these plates is shown in Fig. 2. The photodensitometer and emulsion were calibrated in the following manner: alpha particles from a Po^{208} source were allowed to impinge upon an emulsion after passing through a thin slit. The length of exposure for a number of lines was varied so that a group of lines with known relative intensities was obtained. As a further check the relative intensities for several plates were determined by counting the number of tracks per unit area with a microscope under 97×10 magnification. The values obtained with the microscope agreed with the densitometer readings to within 0.02 for all charge groups.

RESULTS

The results of these measurements are presented in Fig. 3 where the fraction of the total number of particles in each charge group is plotted as a function of

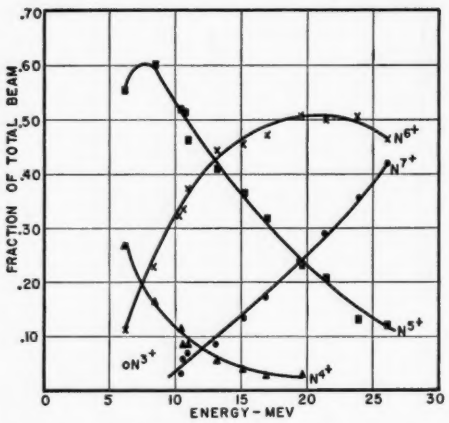


FIG. 3. Equilibrium charge distribution as a function of energy for nitrogen ions in Formvar.

energy from 6 to 26 Mev. Each point on the curve results from the average of several exposures. The probable error in each of the fractions is estimated to be 0.02. The largest part of the error is in the calibration of the photodensitometer. The energies are accurate to better than 0.5 Mev. In the low energy portion of the curves, perhaps below 10 Mev, the energy scale may be distorted to some extent because of straggling, since a large part of the initial energy was lost in the first foil. Also, the initial beam spread of about 600 kev becomes serious at this energy. The range energy relations in nickel were taken from a previous paper.[7] The initial energy of the beam with no absorber present was determined by bombarding a carbon plate directly over slit 1. The relative activity of F^{18} and Na^{24} induced in the carbon was used in the energy determination; the method has been described previously.[12]

In order to insure that an equilibrium charge was being reached in the 50 $\mu g/cm^2$ foil, charge distributions were studied for Formvar foils of 100, 50, 20, and 6 $\mu g/cm^2$ thickness at an energy of 26 Mev. The same charge distribution was obtained for all but the 6 $\mu g/cm^2$ foil, indicating that the charge distribution had reached equilibrium after passing through foils of 20 $\mu g/cm^2$, or more. The equilibrium distribution for 26-Mev ions which passed through an aluminum foil 0.18 mg/cm^2 thick was found to be identical with that observed for Formvar foils. It is to be expected that the Z dependence of the equilibrium distribution is not very great. In any case, the possibility of surface films on the thin films could mask the Z dependence. For these reasons and also because sufficiently thin foils of higher Z were not available, the question of Z dependence was not pursued.

The charge distribution found at 15 Mev does not agree with the results of Stephens and Walker.[8] The distribution which they give would be in agreement with Fig. 3 at 12.5 Mev. As a further check on the energy, we have measured the nitrogen beam energy at the 15-Mev point and at several other points independently of the range-energy curve for nitrogen in nickel. This was accomplished by measuring the range in emulsion of recoil protons from nitrogen-hydrogen collisions in the Formvar at zero degrees to the nitrogen beam. The energy of the protons specifies the energy of the incident nitrogen ions. The energies measured in this way agreed with the energy inferred from the nitrogen range-energy curve within 0.2 Mev.

DISCUSSION

Since the loss-to-capture cross-section ratio increases very rapidly with decreasing binding energy, it is assumed that the captured electrons in the ion will be in the most tightly bound states.[13] Therefore in the ensuing discussion the first electron captured by a completely stripped ion is considered to be the most tightly bound K electron, the second electron, the loosely bound K electron, etc.

At 26 Mev, the main process involved is the capture and loss of the first K electron, while at 6 Mev the main process is capture and loss of the third electron, the most tightly bound L-shell electron. The equations describing the equilibrium state can be written as follows. Let σ_{ij} be the cross section for the change of the ion charge from i to j, while N_k is the fraction of the ions with charge k. The possibility of events in which two electrons are transferred in a single encounter is neglected.[14] Therefore,

$$\begin{aligned} N_7\sigma_{76}&=N_6\sigma_{67},\\ N_6(\sigma_{67}+\sigma_{65})&=N_7\sigma_{76}+N_5\sigma_{56}, \qquad (1)\\ N_5(\sigma_{56}+\sigma_{54})&=N_6\sigma_{65}+N_4\sigma_{45}, \text{ etc.} \end{aligned}$$

with the condition that

$$\sum_{k=0}^{7} N_k=1.$$

The values of the cross sections cannot be determined without a measurement for a nonequilibrium state. One can, however, determine the ratio of loss to capture cross sections for a particular electron from the equilibrium relations since, from Eq. (1):

$$\sigma_{67}/\sigma_{76}=N_7/N_6, \qquad (2)$$

$$\sigma_{56}/\sigma_{65}=N_6/N_5, \qquad (3)$$

$$\sigma_{45}/\sigma_{54}=N_5/N_4. \qquad (4)$$

The ratios of loss to capture cross sections are plotted, in Fig. 4, as a function of energy. The points shown are taken from the measured values in Fig. 3. Bohr[5] predicts that the ratio of the loss to capture cross

[12] H. L. Reynolds and A. Zucker, Phys. Rev. 96, 393 (1954).

[13] N. Bohr and J. Lindhard, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd. 28, No. 7 (1954).

[14] In private communication, R. L. Gluckstern indicates that double electron transfer may be a probable occurrence.

section should be proportional to V^m, where V is the particle velocity and m is a function of the charge of the particle and the material through which it passes. He predicts that for alpha particles in light materials m should be approximately five, while for fission fragments m should be three. From Fig. 4 it can be seen that the value of m for both nitrogen K electrons is about 4.0 and constant over the energy range studied. The value of m for the first L-shell electron is 2.9. Dissanaike[15] has found that for σ_{12}/σ_{21} with alpha particles, m is 5.1 in a solid, while for σ_{01}/σ_{10}, m is 3.08.

For nitrogen ions the value of $(\sigma_{56}/\sigma_{65})/(\sigma_{67}/\sigma_{76})$ is about 4.4. If the two K electrons were considered to have the same bound state velocity, and the electrons in the medium were free, this ratio would be 8, due to the statistics of the process. The cross section σ_{76} would be four times that of σ_{65} since the second electron would have to be captured in a singlet state. The cross section σ_{56} would be twice σ_{67} since, in the former case, two electrons of equal average velocity are available for loss. It is readily apparent that the process is more complicated than this picture.

TABLE I. Values of γ for H and He in metals and for N^{14} in Formvar.

	γ_1[a]	γ_2[b]
H	1.05	...
He	0.78	1.07
N	0.78	1.03

[a] γ_1 refers to the first bound K electron.
[b] γ_2 refers to the second K electron.

The energies at which $N_6=N_7$ and $N_5=N_6$ correspond to the velocities at which $\sigma_{67}=\sigma_{76}$ and $\sigma_{56}=\sigma_{65}$. It has been assumed by Bohr[5] and others that the capture and loss cross sections are nearly equal when the velocity of the ion (V) is equal to the velocity, V_e, of the bound electron in question, or when $\gamma=V_e/V=1$. Knipp and Teller[3] have defined V_e as the electron velocity obtained from a Fermi-Thomas calculation. In a previous paper[7] where the average charge was measured as a function of energy, this definition was used to compare the data with the Knipp and Teller calculations and γ was found to be 0.65. However, it has been the general practice to define V_e as the classical electron velocity obtained from the experimental ionization potential. In order to compare our data with the recent hydrogen and helium data we will use this latter definition since the computations are much simplified. Also, the Fermi-Thomas model is not very significant

[15] G. A. Dissanaike, Phil. Mag. 44, 1051 (1953).

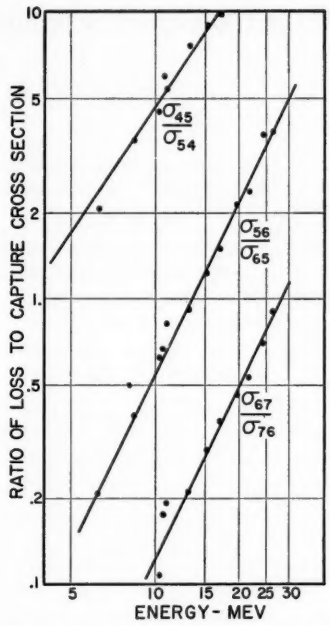


FIG. 4. Electron loss-to-capture cross-section ratios as a function of energy for nitrogen ions in Formvar.

for K electrons. We have used the ionization potentials given by the National Bureau of Standards.[16] Table I contains the values of γ for H, He, and N. The values of γ for H and He are for passage through a metal, while for the nitrogen case the medium is Formvar. The hydrogen value in metal was obtained from Hall,[17] while the helium values in metal are from Dissanaike.[15] It would appear from the similarities for the values for He and N that these values could be applied to other light elements for calculations involving electron capture in the K-shell. Snitzer[18] finds a different value of γ for He ions traversing gases, as is to be expected from Bell's calculation.[6] In general, the average charge of moving ions should be lower in a gas than in a solid.

The average equilibrium charge as a function of energy can be computed from Fig. 3. The average charge curve thus obtained lies about 3 percent higher than that given by Reynolds *et al.*[7] which is within the errors of the two measurements.

We wish to thank Dr. R. S. Livingston for his encouragement in the course of this investigation and Mrs. W. Bowelle for her assistance in scanning.

[16] National Bureau of Standards Circular 467.
[17] T. Hall, Phys. Rev. 79, 504 (1950).
[18] E. Snitzer, Phys. Rev. 89, 1237 (1953).

Orientation and Alignment of Sodium Atoms by Means of Polarized Resonance Radiation*

WILLIAM BRUCE HAWKINS†
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received December 17, 1954)

Sodium atoms have been oriented by means of circularly polarized resonance radiation. The amount of orientation agrees with that calculated from the intensity of the light source used. The polarization of the light scattered from the sodium sample varies in the expected manner when the light intensity and applied magnetic field are varied. Illumination with unpolarized resonance radiation is shown to result in alignment of the sodium atoms. The largest degree of orientation achieved corresponds to an average value for the nuclear, electronic, and total angular momenta of $\bar{M}_I=0.180=0.120I$, $\bar{M}_S=0.035=0.070S$, $\bar{M}_F=0.216=0.108F$, respectively.

INTRODUCTION

KASTLER[1] has proposed that atoms could be oriented by means of circularly polarized resonance radiation. I shall use the word orientation to denote a state in which an ensemble of atoms has an average atomic or nuclear magnetic moment not equal to zero; I prefer this to the more commonly used word polarization in order to avoid confusion with light polarization. I shall retain the common usage of the word alignment to denote a condition in which, while the ensemble average magnetic moment is zero, atomic states with $|M_F|=F$ are more probable than those with $M_F=0$. A substantial quantity of oriented atoms could be used as a source of oriented nuclei for such experiments as the study of angular distribution of radioactive decay products, as a source of polarized electrons, or as a means of enhancing the signal in nuclear magnetic resonance experiments.

Bitter and Brossel[2] have unsuccessfully attempted to induce orientation in mercury by this method. More recently, Brossel, Kastler, and Winter[3] have achieved a positive result with sodium, while Bitter, Lacey, and Richter[4] also report failure with mercury. The present experiment using sodium has already been briefly described.[5,6] Brossel, Cagnac, and Kastler[7] have induced radio-frequency transitions between the Zeeman sublevels of oriented sodium atoms. They observe transitions involving two rf photons corresponding to a change $\Delta M_F=2$ in the axial angular momentum as well as the usual single-photon $\Delta M_F=1$ transitions. An excellent review of the whole subject has been given by Kastler in his recent Holweck lecture.[8]

* This research was supported by the U. S. Atomic Energy Commission and the Higgins Scientific Trust Fund. It has also been described in a dissertation submitted to Princeton University in partial fulfillment of the requirements for the PhD degree.
† National Science Foundation Predoctoral Fellow 1953–54. Present address: Physikalisches Institut der E. T. H., Gloriastrasse 35, Zürich, Switzerland.
[1] A. Kastler, J. phys. et radium **11**, 255 (1950).
[2] F. Bitter and J. Brossel, Phys. Rev. **85**, 1051 (1952).
[3] Brossel, Kastler, and Winter, J. phys. et radium **13**, 668 (1952).
[4] Bitter, Lacey, and Richter, Revs. Modern Phys. **25**, 174 (1953).
[5] W. B. Hawkins and R. H. Dicke, Phys. Rev. **91**, 1008 (1953).
[6] W. B. Hawkins, Phys. Rev. **96**, 532 (1954).
[7] Brossel, Cagnac, and Kastler, J. phys. et radium **15**, 6 (1954).
[8] A. Kastler, Proc. Phys. Soc. (London) **A67**, 853 (1954).

THEORY OF ORIENTATION

General

The simplest atomic system which can be oriented by means of circularly polarized resonance radiation is one having a ground state of total angular momentum $F=\frac{1}{2}$ and an optical first excited state having $F=\frac{3}{2}$. If such an atom is in the ground state sublevel with axial angular momentum $M_F=+\frac{1}{2}$ and absorbs a circularly polarized photon with axial angular momentum $+1$, it will go into the $M_F=+\frac{3}{2}$ substate of the excited level. By virtue of the electric dipole selection rules which are assumed for the transition, this excited atom must return to the substate $M_F=+\frac{1}{2}$ when it radiates. An atom originally in the $M_F=-\frac{1}{2}$ ground state absorbing such a photon can return to either ground state sublevel when it radiates. Thus the result of the scattering of this kind of light by an ensemble of such atoms is that all atoms originally in the $M_F=+\frac{1}{2}$ state will return there, while some of those with $M_F=-\frac{1}{2}$ will be transferred to the $M_F=+\frac{1}{2}$ state by the scattering process. The ensemble has been partially oriented.

More complicated atoms having larger angular momenta or hyperfine structure behave in essentially the same way, the ground-state sublevel of largest (or smallest if the sense of the light polarization be reversed) magnetic quantum number gaining atoms at the expense of the others. Figure 1 shows how the orientation process works for one of the hyperfine transitions in Na^{23} together with the complete energy

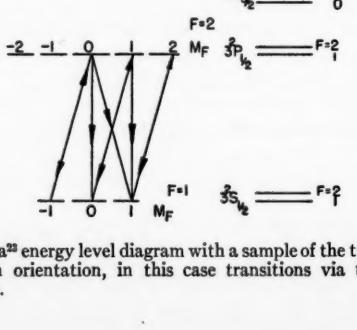


FIG. 1. Na^{23} energy level diagram with a sample of the transitions involved in orientation, in this case transitions via the $3\,^2P_{\frac{1}{2}}$, $F=2$ states.

level diagram. Although the magnetic quantum number can change by at most two units during a single scattering process, the scattering of further photons will increase the orientation if the partially oriented ensemble remains undisturbed. However, its members may be disturbed between photon-scattering events by collisions with other atoms or by the presence of magnetic fields.

A magnetic field perpendicular to the axis of quantization defined by the direction of the incident light will cause a mixing of all the M_F states for a particular F. Although the resulting state is not one of thermal equilibrium, it may be treated as such for present purposes. A field of about 10^{-3} gauss is sufficient for thorough mixing in a time of the order of 10^{-4} second, the time of observation in this experiment. However, the situation is quite different if in addition to this transverse component a larger axial field component is present, for a field nearly parallel to the axis of quantization mixes the states only slightly. The angle between the axis and whatever magnetic field is present is more important than the size of the field, provided that it is not large enough to seriously alter the atomic wave functions. The fields normally applied in this experiment did not appear to be large enough to cause significant alteration. Such changes in wave functions, while rendering the calculations for zero field invalid, do not interfere with the basis process of "optical pumping." Their avoidance is thus a matter of convenience in the interpretation rather than an essential of the experiment.

Collisions with other atoms of suitable kinds can result in disorientation either by electron exchange or by mutual spin-flipping. This source of disturbance can be eliminated by keeping the sodium mean free path large compared to the dimensions of the apparatus, which is equivalent in this experiment to keeping the pressure below 10^{-2} mm of Hg. The normal pressure was 10^{-6} to 10^{-4} mm of Hg.

Sodium Resonance Fluorescence

The electronic ground state of sodium is a $3\,^2S_{\frac{1}{2}}$ state which combines with the nuclear spin of the stable isotope, $I=\frac{3}{2}$, in two hyperfine states having $F=1$ and 2. The first excited state is a doublet having term values $3\,^2P_{\frac{1}{2}}$ and $3\,^2P_{\frac{3}{2}}$, the lower of which ($J=\frac{1}{2}$) has the same hyperfine structure as the ground state, while the upper has four hyperfine states with F ranging from 0 to 3 (see Fig. 1).

Since the smallest hyperfine splitting is only 2.6 times the natural level breadth, the treatment of resonance fluorescence given in Heitler[9] must be extended to include the effect of the overlapping of levels. The derivation is very similar to Heitler's, and only the result will be given. The probability of an atom being in a given final state (at $t=\infty$ when absorption is known to have taken place) is

$$|b_m(\infty)|^2=\frac{4\pi^2\rho_\omega^2}{\Gamma\hbar^4}\sum_{n,n'}\frac{H_{mn}H_{n0}H_{mn'}{}^*H_{n'0}{}^*}{\gamma-i\omega_{nn'}}, \quad (1)$$

where ρ_ω is the density of radiation oscillators at the resonance frequency, $1/\Gamma$ is the lifetime of the initial state, $1/\gamma$ that of the intermediate state, and $\omega_{nn'}$ the energy difference between the two indicated intermediate states. The index 0 refers to the initial state, n and n' to intermediate states, and m to the final state. This result has been averaged over the frequency distributions of the emitted and absorbed photons, assuming a continuous spectrum incident. This assumption is justified for a light source emitting lines much broader than the absorption Doppler breadth such as the one used in this experiment.

The matrix elements H_{mn} and H_{n0} for the emission and absorption processes may be obtained from the formulas in Condon and Shortley,[10] page 63, although a unitary transformation must be applied to obtain the values in the desired representation having the quantum numbers I, J, F, and M_F. The matrix of this unitary transformation may be constructed from the tables in Condon and Shortley, page 76. Applying this transformation, one obtains the matrix elements of electron momentum in the desired representation. The elements H_{mn} and H_{n0} of the Hamiltonian are derived from these by taking the inner product of the electron momentum and the vector potential of the radiation field in the dipole wave approximation, care being taken to appropriately restrict the polarization of the incident wave.

It may be seen from Eq. (1) that a pair of intermediate levels will contribute an interference term whose ratio to the direct transition terms is $2/(\omega_{nn'}{}^2+\gamma^2)$. The natural line breadth γ is[11] 6.1×10^7 radians/sec, while the total hyperfine splittings used in the calculations are [12] 1.5×10^9 radians/sec for the $3\,^2P_{\frac{3}{2}}$ state and 9.4×10^8 radians/sec for the $3\,^2P_{\frac{1}{2}}$ state. Applying the Landé interval rule, it is found that only the levels $F=0$ and $F=1$ of the latter state ($\omega_{01}=2.6\gamma$) are close enough to contribute interference terms as large as 10 percent of the direct transition terms. More recent measurements[13] give a smaller total splitting in the $3\,^2P_{\frac{1}{2}}$ state, as well as a small deviation from the interval rule, but the difference does not significantly affect the results of the present calculations.

The transition probabilities also depend on the intensity ratio in the light source of the two fine structure components of the sodium D-line. The source used yielded equal peak intensities of the two

[9] W. Heitler, *The Quantum Theory of Radiation* (Oxford University Press, London, 1944), second edition, pp. 137–142.

[10] E. U. Condon and G. H. Shortley, *Theory of Atomic Spectra* (Cambridge University Press, London, 1951).

[11] L. Larrick, Phys. Rev. **46**, 581 (1934).

[12] L. P. Granath and C. M. Van Atta, Phys. Rev. **44**, 935 (1933).

[13] P. Sagalyn, Phys. Rev. **94**, 885 (1954).

TABLE I. Probabilities of transitions induced by resonance radiation between substates of the sodium $3\,^2P_{\frac{1}{2}}$ ground state complex for equal incident fluxes of the two D-line components. Column one gives the quantum numbers F and M_F of the initial states, labelling the rows; and the other columns are headed by the quantum numbers F and M_F of the final states. The probabilities are so normalized that each initial state has a total absorption probability of one.

Initial F, M_F	Final F, M_F 1,−1	1,0	1,1	2,−2	2,−1	2,0	2,1	2,2
1,−1	0.343	0.167	0.268	0	0.085	0.056	0.082	0
1, 0	0	0.312	0.299	0	0	0.140	0.085	0.164
1, 1	0	0	0.500	0	0	0	0.167	0.333
2,−2	0.169	0.164	0	0.321	0.193	0.154	0	0
2,−1	0.085	0.167	0.082	0	0.229	0.207	0.229	0
2, 0	0	0.140	0.138	0	0	0.292	0.276	0.154
2, 1	0	0	0.167	0	0	0	0.500	0.333
2, 2	0	0	0	0	0	0	0	1.000

lines; and therefore, because of the difference in statistical weights of the two levels, the total absorption probability leading to the $3\,^2P_{\frac{3}{2}}$ state is twice that for the $3\,^2P_{\frac{1}{2}}$ level. This having been taken into account, Table I gives the total relative probabilities of resonance fluorescence between given initial and final states via all possible intermediate states. The columns are labeled by the quantum numbers of the final states, and the rows by those of the initial states. Both initial and final states are all members of the $3\,^2S_{\frac{1}{2}}$ ground state complex. The total absorption probabilities of all the initial states are the same, which is only true for equal intensities in the source of the two components of the doublet.

Table I may be regarded as a matrix operator. The vector on which it operates has components proportional to the occupation probabilities of the various initial states for whatever ensemble of atoms is under consideration. Similarly, the vector resulting from the operation gives the occupation probabilities for the same ensemble after each of its members has scattered one and only one photon. The initial ensemble is normally taken as being in thermal equilibrium, and by n applications of the operator one may find the occupation probabilities of an ensemble each of whose members has scattered exactly n photons.

One may simplify the computation of the polarization of the scattered light by observing that each element in the matrix operator just described corresponds to a definite polarization component, σ or π, in the scattered radiation. Summing the elements of like polarization in each row of the matrix produces a pair of vectors whose components are labeled by the quantum numbers of initial states. The inner product of one of these vectors with the occupation vector of an ensemble is proportional to the amount of the corresponding polarization emitted when each member of that ensemble scatters one photon. For example, applying this process to an ensemble whose members have each scattered two photons, one can compute the polarization emitted when a third photon is scattered.

Table II gives the state occupation probability distributions of ensembles whose members have each scattered a definite number, n, of photons. It also gives the expectation values for these ensembles of the observables M_F, M_I, and M_S, the axial projections of the total, nuclear, and electronic angular momenta, as well as the polarization ratio $R_n=(\sigma-\pi)/(\sigma+\pi)$ of the light scattered at 90° to the axis of quantization. π radiation has its electric vector parallel to the axis, σ, perpendicular, both being plane polarized at this scattering angle.

Of course, no physically realizable ensemble is composed of members all of which have scattered the same number of photons. However, the expectation value of an observable for a physical ensemble may be computed by taking a weighted average of the expectation values for the several such idealized ensembles. The value of the observable for the ith ensemble is weighted by the probability P_i of an atom in the physical ensemble scattering i photons.

To compute P_i, consider a beam of sodium atoms passing through an illuminated region such that each atom scatters on the average N photons in the L cm of illuminated beam. The probability of absorbing a photon in any interval of path dl is $(N/L)dl$, and the probability of not absorbing any photon at all is

$$P_0=(1-Ndl/L)^{L/dl}=e^{-N}, \tag{2}$$

of absorbing one and only one photon is

$$P_1=\int_0^L \frac{N}{L}e^{-N}dl=Ne^{-N}, \tag{3}$$

and of absorbing exactly i photons is

$$P_i=N^ie^{-N}/i!. \tag{4}$$

EXPERIMENTAL PROCEDURE

The apparatus consisted of a source of sodium resonance radiation, a sodium sample in the form of a

TABLE II. Expectation values of various quantities for ensembles each of whose members has scattered n photons. R_n is the polarization ratio of photons scattered at 90° for the nth scattering process.

	Occupation numbers											
		$F=1$			$F=2$							
n	$M_F=1$	0	−1	2	1	0	−1	−2	M_F	M_I	M_S	R_n
1	0.182	0.119	0.075	0.247	0.168	0.106	0.063	0.040	0.626	0.523	0.103	0.070
2	0.194	0.082	0.038	0.399	0.174	0.071	0.029	0.012	1.075	0.884	0.191	0.173
3	0.173	0.049	0.018	0.546	0.156	0.042	0.012	0.004	1.384	1.116	0.268	0.313

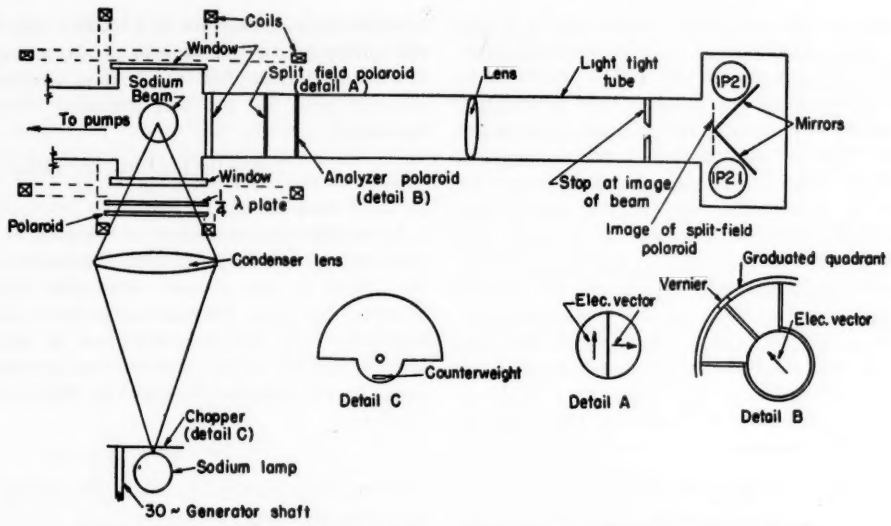


FIG. 2. Optical system used to orient sodium and to detect the orientation when produced.

beam, coils to control the magnetic field seen by the sample, and equipment to detect the polarization of the light scattered by the sample. A view of the optical system is shown in Fig. 2.

A Phillips sodium spectral lamp No. 93122E was used, together with a sheet of polaroid and a quarter wave plate, to produce circularly polarized sodium resonance radiation. A lens formed an image of the lamp at the beam position, and the light was interrupted at 30 cycles per second by a semicircular metal segment on the shaft of a small motor generator. The 30-cycle/sec output from the generator was used as the reference signal for the lock-in amplifier that formed part of the polarization detector.

A sodium beam was used as a simple means of insuring relative freedom from foreign gas contamination. The beam source was a stainless steel oven heated to 100°–200°C by 60-cycle/sec ac. The beam was collimated into a cone of 9° half-angle and caught by a liquid nitrogen trap after leaving the illuminated region. The beam chamber was evacuated by an oil diffusion pump provided with a solid CO_2 cooled trap to intercept oil vapor. With both traps cold, a vacuum of between 10^{-4} and 10^{-6} mm of Hg could be obtained, depending on the beam source conditions.

In the early stages of the experiment, the magnetic field applied was quite inhomogeneous, especially the axial component (that parallel to the direction of the incident light), since it was supplied by two solenoids separated axially by a distance equal to their diameter, the diameter being no larger than the region in which the sodium was observed. When it was found that this inhomogeneity increased the difficulty of interpreting the data, the first set of coils was replaced by three mutually perpendicular Helmholtz pairs, each producing a field of 4 gauss/ampere within the accuracy of mounting them to fulfill the Helmholtz condition. No refined technique of mounting them was used, neither the field calibration nor the homogeneity being sufficiently critical to the experiment to require it. The smallest of the three pairs had a radius of 11.4 cm and the largest 19.0, while the length of sodium beam observed was 3.8 cm.

The polarization ratio $(\sigma-\pi)/(\sigma+\pi)$ of the light scattered by the sodium beam was determined by an optical bridge whose parts are indicated in Fig. 2. All observations were made at a scattering angle of 90°. The light passed through a polaroid filter consisting of two semicircular segments each transmitting only σ or only π radiation. At 90° to the axis of quantization, each component is plane polarized with their planes of polarization perpendicular to each other. Each half of this split-field polaroid was imaged by a lens on one of two 1P21 photomultiplier tubes. Between the split-field polaroid and the lens was a rotatable polaroid (hereafter called the analyzer) provided with a graduated quadrant and vernier. Its function was to produce known changes in the intensity ratio of the two light polarization components and thus calibrate the detector. It was normally set at 45° in order to transmit equal amounts of the two components. All this equipment was enclosed in a light-tight tube. To reduce stray light to a minimum, the detector was placed opposite the vacuum pumping lead; and a stop was placed at the image of the beam (see Fig. 2) to intercept light scattered from the edges of the observation window.

Since the incident light was modulated at 30 cycles/sec, ac detection could be used. The output of one 1P21 was inverted in phase and added to that of the other, the combined signal being fed into a 30-cycle/sec lock-in amplifier. The output of this amplifier appeared on an Esterline-Angus recording milliammeter. The

system thus formed an optical bridge giving a null signal for equal outputs from the two photomultiplier tubes. Since the gains of the two 1P21's could not be depended on to be the same, absolute light polarization ratios could not be measured. All polarization measurements were therefore referred to a standard condition, that in which sodium orientation was destroyed by application of a magnetic field perpendicular to the axis. The field of the earth was found to serve this purpose excellently. With unoriented sodium, the bridge was approximately balanced by adjusting the applied voltage, and thus the gain, of one photomultiplier. The difference between the resulting signal and that for any other sodium sample condition was reduced to a light polarization difference by comparing it with the signal change produced by rotating the analyzer polaroid a known amount.

RESULTS AND DISCUSSION

The sodium lamp used was compared with a white lamp of known intensity by means of visual observation with a flicker photometer (average of 92 comparisons) and found to emit 2.72×10^{16} photons/(cm^2 sec steradian) perpendicular to its surface. If an image of this source is formed by a lens, it can be shown that the photon flux through the image is simply $I\Omega$, where I is the surface intensity of the source as given above and Ω is the solid angle subtended by the lens at the image. For a given lens, the maximum solid angle is determined by the requirement that the image of the light source be at least as big as the region occupied by the sodium beam. In practice, the solid angle was limited not by the lens, but by the window admitting light to the vacuum system, which subtended an angle at the sodium beam of $\pi/16$ steradian; since the polaroid had a transmission of 0.365, the total photon flux at the beam was therefore 1.95×10^{15} photons/cm^2 sec.

To determine the intensity ratio and line widths of the two fine structure components, the light source was photographed through a Lummer-Gehrcke plate (crossed with a prism), for the loan of which I am indebted to Professor A. G. Shenstone. The density of the photographic images was measured with a recording densitometer owned by the R. C. A. laboratory in Penn's Neck, New Jersey. I am grateful to Dr. Hillier of that laboratory for permission to use the instrument and to Dr. H. Halma for making the measurements with it. The two components of the doublet were found to have approximately the same peak intensities; their widths were 10^{10} cycles/sec for the component radiated by atoms in the $3\,^2P_{\frac{1}{2}}$ level and 1.4×10^{10} cycles/sec for the $3\,^2P_{\frac{3}{2}}$ component. This equality of peak intensities, together with the extreme breadth (20 times Doppler breadth), suggests that the source is so self-reversed as to be effectively a blackbody at the center of the lines. This is quite reasonable since we shall see that approximately 10^{10} atoms cm^{-2} completely absorb all light at the resonance frequency. For a thickness of 1 cm, this corresponds to a pressure of 2.5×10^{-7} mm of Hg, and the sodium partial pressure in the lamp was certainly much higher than this. From these observations, the common peak flux per unit frequency at the sodium beam is

$$I_0=1.29\times10^4 \text{ photons/cm}^2 \text{ radian}$$

for each component.

In sodium the transition to the first excited state complex accounts for 0.975 of all the radiation absorbed by atoms in the ground state (see Condon and Shortley,[10] p. 149). For present purposes the oscillator strength of the line may therefore be taken as one, and the number of photons absorbed per atom is thus given by the formula (Heitler,[9] p. 108) for a classical oscillator,

$$n=2\pi^2 r_0 c I_0, \tag{5}$$

where r_0 is the classical radius of the electron. Thus for the given light flux,

$$n=2.14\times10^8 \text{ photons/sec atom.}$$

A sodium beam source temperature of 200°C corresponds to a velocity of 4.8×10^4 cm/sec. Since the illuminated length of beam was 3.8 cm, each atom might be expected to absorb on the average 0.17 photon.

Clearly the number of sodium atoms which can be oriented is no larger than that number which absorbs all the incident light. A larger sample will result in multiple scattering of photons and a lower orientation per atom. For a sodium sample (not necessarily a beam) at a temperature in the neighborhood of 200°C, the Doppler breadth of the absorbed line is approximately 2×10^9 radians/sec. The total number of photons which can be absorbed is roughly equal to the line breadth times the peak intensity, or 2.4×10^{13} photons/cm^2 sec. Since each atom absorbs 2.1×10^8 photons/sec, 10^{10} atoms/cm^2 will absorb all photons in the usable range. Other considerations, based on the tendency of a gas to radiate coherently,[14] thus reradiating photons of the same polarization as those absorbed, limit the sample to about one-fifth this number.

Although the fact is obscured by the method of derivation, the sample size is independent of the light intensity; for as the intensity increases, each atom absorbs more photons. In other words, the absorption cross section is independent of the light intensity.

This limitation of the sample size is naturally a serious limitation on any experiments it is desired to perform with atoms or nuclei oriented by this technique. As an example, if it were desired to study the angular distribution of radioactive emissions, an isotope with a half-life of one year would produce 130 counts/min in a perfectly efficient counter subtending a solid angle of $4\pi\times10^{-2}$ steradian if the radiation were isotropic. Again, the hyperfine structure resonance of

[14] R. H. Dicke, Phys. Rev. **93**, 99 (1954).

Na^{23} could be observed with a signal-to-noise ratio of about 500 assuming an orientation of 10 percent. The signal-to-noise ratio varies as the square of the amount of orientation.

It was found that the sodium source temperature (and beam density) could be allowed to vary over wide limits without changing the experimental results. This is interpreted to mean that the density was at all times well below the allowable maximum. No attempt was made to observe the actual source temperature beyond rough estimates.

Since the rate of light absorption is now known, the probability of an atom scattering any given number of photons i may be computed from Eq. (4). The average polarization of the light scattered is given by

$$\bar{R}=\sum_{i=1}^{\infty} R_i \sum_{j=i}^{\infty} P_j \Big/ \sum_{i=1}^{\infty} iP_i, \qquad (6)$$

where the R_i are taken from Table II. For the given absorption probability, P_3 is only 5 percent of P_2, and all higher P_i's may be neglected. The average polarization is then

$$\bar{R}=0.0789.$$

When orientation is destroyed by application of a transverse magnetic field, the polarization of the scattered light is not quite equal to R_1, as one would expect if true thermal equilibrium were produced. It can be shown that a slight degree of *alignment* remains and that the average scattered polarization in this case is

$$\bar{R}=0.0710,$$

while the value of R_1 is 0.0700. The change in light polarization expected upon destruction of orientation by a transverse field is therefore

$$\Delta\bar{R}_{\text{th}}=0.0079.$$

Twenty-one observations of this change were made over a period of five months, and their average was

$$\Delta\bar{R}_{\text{obs}}=0.0097.$$

The largest and smallest values observed were

$$\Delta\bar{R}_{\text{min}}=0.0058,$$

$$\Delta\bar{R}_{\text{max}}=0.0129.$$

Since the intensity of illumination may easily be affected by changes in position of the light source, line voltage changes, and other effects not under experimental control, this variation is not considered alarming.

The agreement with the predicted value of ΔR is considered good in view of the uncertainties involved in the use of the flicker photometer and the photographic measurement of the line width, especially since the latter was not done with as much skill as might be desirable.

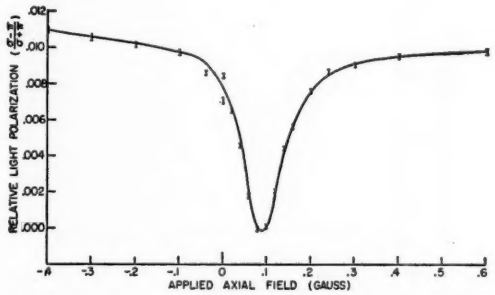


FIG. 3. Polarization ratio of the scattered resonance radiation as a function of the applied axial field component, incident light circularly polarized.

Although the transverse components of the ambient magnetic field were compensated as well as possible, one could not hope to reduce them below the 10^{-3} gauss necessary to avoid disturbing the sodium orientation. In addition, no attempt was made to compensate ac fields. Thus one might expect that orientation would disappear for insufficiently large values of the axial field. This was indeed the case, and a representative curve of light polarization *versus* applied axial field is shown in Fig. 3. The minimum is displaced from zero by the presence of a terrestrial component of 0.1 gauss in the axial direction. It should be stressed that ideally the incident light is sufficient to define the axis of quantization, and it is only necessary to apply a field to overcome the disturbing effects of fields already present.

When the measurements are extended to axial fields of several gauss, the polarization at first increases for one field direction and decreases for the other as the absolute value of the field is increased. In both directions it eventually goes through a minimum and then increases steadily. This is believed to be the result of the alteration of atomic wave functions by the magnetic field. The asymmetry is a result of the asymmetry of the light polarization and disappears for unpolarized light although the changes in polarization with increasing field remain.

The measurements of $\Delta\bar{R}$ discussed above were all taken from curves similar to Fig. 3. at about -0.2 gauss applied where the axial field is sufficiently large to allow orientation to proceed undisturbed but not large enough to change the wave functions appreciably.

The vacuum system was later modified to increase by a factor of 2.5 the illuminated solid angle seen by the sodium beam. The change in light polarization upon destruction of orientation then became $\Delta\bar{R}=0.0175$, an increase by a factor of 1.8. This increase could reasonably be expected to be less than the increase in solid angle because the additional light is all moving at an angle to the axis and is therefore less effective in producing orientation.

This polarization change of 1.75 percent corresponds theoretically to an effective average of 0.365 photon

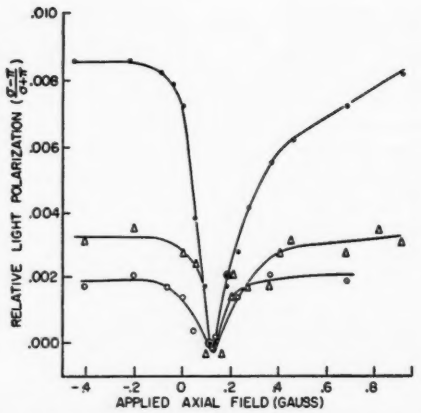


FIG. 4. Polarization ratio of the scattered resonance radiation as a function of incident intensity of circularly polarized radiation. Incident light intensities in ratio 1.00:0.39:0.17 for upper, middle, and lower curves, respectively.

scattered by each atom. For this probability of photon scattering, the ensemble averages of the expectation values of nuclear, total, and electronic angular momenta are, respectively,

$$\bar{M}_I = 0.180 = 0.120I,$$
$$\bar{M}_F = 0.216 = 0.108F,$$
$$\bar{M}_S = 0.035 = 0.070S.$$

In the case of F, the fractional figure is given relative to the larger F value of the two hyperfine states, for if orientation were complete, all atoms would be in the state $F=2$.

In order to obtain an experimental check on the polarization of the first photon scattered by an unoriented atom, the absolute polarization of light scattered by sodium atoms in the earth's magnetic field was measured. An attempt was made to do this by using the electronic detector already described and a modified scheme of analyzing polaroids, but the results were not self-consistent. Visual methods were therefore used. A Babinet compensator followed by a Nichol prism shows dark bands crossing the field of view if the polarization ratio of the light incident on the arrangement is greater than about 0.01. A glass plate between the light source and the Babinet compensator may be used to introduce an opposite polarization by rotating it in a plane containing the electric vector of the observed polarization. The polarization of the light may be calculated from the angle at which the bands disappear.[15] The angle observed was $34.2\pm1.5°$ (average of ten measurements) when a glass plate of refractive index 1.52 was used. The index was not known exactly and limits of ±0.02 were assigned which account for about one-quarter of the error quoted below for the polarization. The resulting polarization ratio is

$$R = 0.049 \pm 0.008,$$

which differs from the predicted value of 0.071 by 2.8 times the probable error. An error of this size has a probability of 6 percent, viewed as a random deviation.

To explain this result by an error in the measurement of relative intensities of the two components emitted by the source would require the assumption that the $3\,^2P_{\frac{1}{2}}$ component had only one-half the intensity of the other. Not only does this large an error in the photometric measurement seem highly unlikely, but it contradicts the very plausible assumption of blackbody emission at the line centers. An incorrect value for the hyperfine splittings would not affect the polarization by such a large amount, for the predicted polarization is decreased by wider separation of the levels, and with evels completely separated the polarization is only decreased to 0.069. Moreover, the more recent values of the splitting lead to a smaller level separation than the value used.

The intensity of the incident light was varied by means of neutral density filters. By a simple algebraic manipulation of the probabilities P_i of Eq. (4), it can be shown that if P_3 is neglected, the polarization of the scattered light is proportional to $N/(N+1)$ where N is the average number of photons scattered per atom. Since N is small compared to 1 ($N \leqslant 0.2$ in Table III), the polarization can be considered proportional to the light intensity. Figure 4 shows a sample set of curves taken with two filters of different density and without any filter. The asymmetry in these curves was a result of the inhomogeneity of the field produced by the first set of coils used. Table III gives the change in light polarization when orientation is destroyed for two sets of observations similar to those in Fig. 4. The light intensities were determined from the size of the signal produced by rotating the analyzer polaroid a known amount and agree with the filter densities given by the manufacturer. The variation is seen to be linear as expected.

Representative data are given in Fig. 5 and Fig. 6 showing the effect of varying a transverse component of the magnetic field keeping the axial component fixed. The axial components are different for the two figures. The solid curves are theoretical curves computed by assuming that many cycles of Larmor precession occur between successive photon scattering events for any one atom. Since these events are random, only time

TABLE III. Change in scattered light polarization ($\Delta\bar{R}$) when orientation is destroyed as a function of incident light intensity.

Light intensity (arbitrary units)	$\Delta\bar{R}$ (arbitrary units) run 1	run 2
1.00	1.00	1.00
0.39	0.36±0.1	0.38±0.1
0.17	0.20±0.1	0.18±0.1

[15] R. W. Wood, *Physical Optics* (Macmillan Company, New York, 1934), third edition, pp. 341–343.

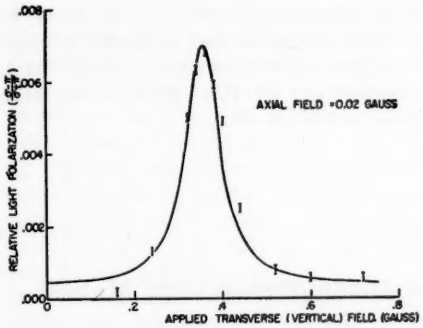


FIG. 5. Polarization ratio of the scattered resonance radiation as a function of one applied transverse field component; the solid curve is computed theoretically.

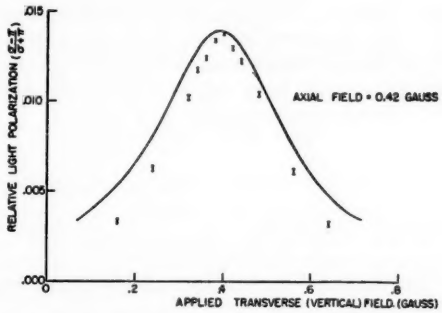


FIG. 6. Polarization ratio of the scattered resonance radiation as a function of one applied transverse field component; the solid curve is computed theoretically.

average results are required. Time-average projection operators of the z-component of angular momentum in a rotating coordinate system were derived. These were used to find the time-average occupation probabilities of the precessing atoms in eigenstates of M_F referred to stationary axes. The calculation is straightforward but too lengthy to be reproduced here. It results in the following expression for the polarization of the second-scattered photon (R_2) as a function of the angle between the magnetic field and the axis of quantization defined by the incident light:

$$R_2(\theta)=\frac{9.8\times10^{-4}\cos^8\theta-0.115\cos^6\theta+4.14\cos^4\theta+4.58\cos^2\theta+7.0}{-10.8\times10^{-4}\cos^8\theta+3.86\times10^{-2}\cos^6\theta-1.28\cos^4\theta-1.50\cos^2\theta+93.4}. \quad (7)$$

This reduces to the correct value of R_2 for $\theta=0$.

The data shown in the figures are typical in that the agreement with theory is good for small axial fields and not for larger ones, the experimental points always falling below the theoretical curve. The only adjustable parameter is the value of the polarization at the maximum of the curve. The disagreement may well be a result of the alteration of atomic wave functions by the larger total fields.

When the incident radiation is unpolarized, the curve of scattered light polarization *versus* axial magnetic field is similar to Fig. 3 in every respect except the light polarization scale. The total change in light polarization is only $\Delta\bar{R}=0.0048$. This may be explained by considering the incident light as composed of equal parts of the two senses of circular polarization. Table II shows that scattering of a circularly polarized photon reduces the population in the $M_F=0$ states independently of its sense of rotation. Consequently, scattering of unpolarized light will increase the populations of the states with $|M_F|=F$ relative to those with $M_F=0$. The atomic ensemble is then in a condition of partial alignment. Among the atoms which scatter two photons there are some which scatter two with the same sense of rotation and contribute to the alignment of the ensemble as well as others which scatter photons of opposite senses and do not contribute significantly. Since the effect of unpolarized resonance radiation on a completely aligned ensemble is to reduce the alignment, the process does not proceed to complete alignment for very large light intensity but rather reaches an equilibrium partial alignment. This possibility of alignment by unpolarized light was briefly mentioned by Kastler in the article already cited.[1]

Table IV gives the probability of an atom being in various angular momentum states after scattering one to four photons, as well as the values at equilibrium ($n=\infty$). These were computed by a straightforward application of Table I considered as a matrix operator. Although the approach to equilibrium is rapid, the equilibrium alignment is small. However, the difference in population of corresponding magnetic substates of the two hyperfine levels is still many times that resulting from the energy separation in thermal equilibrium at room temperature.

Combining the photon polarizations in Table IV with the figure (obtained from the degree of orientation for the same amount of polarized light) of 0.365 photon absorbed per atom, the total change in light polarization expected upon destruction of alignment is calculated to be $\Delta\bar{R}=0.0034$, compared to an observed value of 0.0048. This indicates that the incident light may not have been completely unpolarized.

There is no reason to believe that the effect should be confined to sodium, since the argument for the produc-

TABLE IV. Occupation probabilities for ensembles each of whose members has scattered n photons of unpolarized light, including the equilibrium state ($n=\infty$). P_1 and P_2 are the probabilities of atoms being in states with F equal to 1 and 2. R_n is the polarization ratio of photons scattered at 90° in the nth scattering process.

	$F=1$			$F=2$				
n	$M_F=\pm1$	0	±2	±1	0	P_1	P_2	R_n
1	0.128	0.119	0.144	0.115	0.106	0.375	0.624	0.0700
2	0.128	0.116	0.154	0.110	0.100	0.372	0.628	0.0905
3	0.126	0.115	0.161	0.108	0.099	0.367	0.637	0.0997
4	0.126	0.115	0.162	0.108	0.099	0.367	0.639	0.1049
∞	0.125	0.115	0.163	0.106	0.098	0.365	0.636	···

tion of alignment, like that for orientation, is quite general. The argument applies, however, only when the light is incident from a restricted part of a sphere, for isotropic unpolarized light defines no axis with respect to which alignment can occur.

I wish to express my gratitude to Dr. R. H. Dicke, without whose inspiration and invaluable counsel this research could not have been undertaken or completed. I am also indebted to Dr. D. R. Hamilton for assistance in the early part of the work.

PHYSICAL REVIEW VOLUME 98, NUMBER 2 APRIL 15, 1955

Angular Correlation of Photons from Positron Annihilation in Light Metals*

R. E. GREEN† AND A. T. STEWART‡
Physics Department, Dalhousie University, Halifax, Nova Scotia
(Received November 26, 1954)

The angular correlation of gamma rays from positron annihilation in the light metals, Li, Be, Na, Mg, Al, K, and Ca, has been measured. The deviation of the two photons from colinearity is a measure of the total momentum of the annihilating electron-positron pair. If it is assumed that the positrons are thermalized before annihilation and that they annihilate only with electrons of a Sommerfeld free electron gas, the experimental results may be used to determine the Fermi energy of these electrons. It is interesting that the Fermi energies so determined are in good agreement with calculated values.

I. INTRODUCTION

THE possibility of determining electron velocities in metals from positron annihilation radiation was first pointed out by DuMond *et al.*[1] In their examination of the 0.511-Mev annihilation radiation these authors showed that the measured line width exceeded the instrumental width and ascribed the excess to motion of the electrons with which the positrons annihilate. Since that time DeBenedetti *et al.*[2] and other experimenters[3,4] have measured angular correlation of the two annihilation photons to determine the motion of the center of mass of the annihilating pair. The results of all these experimenters indicate that the center of mass of the annihilating pair has a momentum which is comparable with the momentum of conduction electrons in metals.

This paper reports an experiment in which the angular correlation of the annihilation photons was measured with slightly better angular resolution than had been previously used. Whereas other experimenters have attempted to relate the shape of the "tail" of the angular correlation curve to electron velocity, in this paper it is shown that the width of the observed curve is directly related to the momentum of the annihilating electrons.

Measurements in several metals have been made and the results are compared with theory.

II. EXPERIMENTAL

A schematic diagram of the apparatus is shown in Fig. 1. A source of annihilation radiation was placed at the middle of the horizontal line joining the two scintillation detectors. The rate of coincidences in the two detectors was measured as a function of the vertical displacement of the source. When corrected for background, this coincidence rate curve gives the angular correlation of the gamma-ray pairs.

The positron emitter was made by evaporating an aqueous solution of sodium chloride containing about 0.4 mC of Na^{22} on a thin (2 mg/cm^2) aluminum foil. The approximately 0.5-in. diameter deposit and supporting foil were covered on both sides by another piece of the same foil. A sheet of the specimen metal

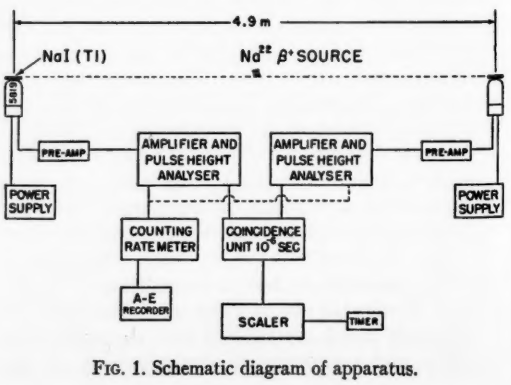


FIG. 1. Schematic diagram of apparatus.

* Part of this work was submitted as a thesis by R. E. Green toward the M Sc Degree at Dalhousie University.

† Now at Physics Department, McGill University, Montreal, Quebec.

‡ Member of the staff of Atomic Energy of Canada Limited, Chalk River, Ontario, on temporary appointment to Dalhousie University during the academic term 1953–54.

[1] DuMond, Lind, and Watson, Phys. Rev. **75**, 1226 (1949).

[2] DeBenedetti, Cowan, Konneker, and Primakoff, Phys. Rev. **77**, 205 (1950).

[3] P. E. Argyle and J. B. Warren, Can. J. Phys. **29**, 32 (1951), and J. B. Warren and G. M. Griffiths, Can. J. Phys. **29**, 325 (1951).

[4] H. Maier-Leibnitz, Z. Naturforsch. **6a**, 663 (1951).

was clamped to each side of the emitter, the resulting sandwich being the source of annihilation radiation. For measurements with the alkali metals and with some others, the emitter was protected by additional (5 mg/cm^2) aluminum foils. The flat, roughly circular, source was tipped to make an angle of about 10° with a line joining the detectors so that the apparent thickness (vertical dimension) of the source of annihilation radiation was independent of positron range in the metal specimen. As seen from the detectors, the source appeared from 3 to 4 mm thick.

The detectors were two NaI(Tl) crystal slabs approximately 2 in.×1½ in.×¼ in. mounted flat on RCA 5819 photomultiplier tubes. The 2 in.×¼ in. face was toward the source.

To eliminate pulses due to photons scattered in the crystal mount and photomultiplier tubes, the pulses from each detector were passed to a pulse height analyzer[5] set to select pulses of amplitude corresponding to a small energy range about 0.51 Mev. Coincidences within a resolving time of one microsecond between selected pulses from the two detectors were counted.

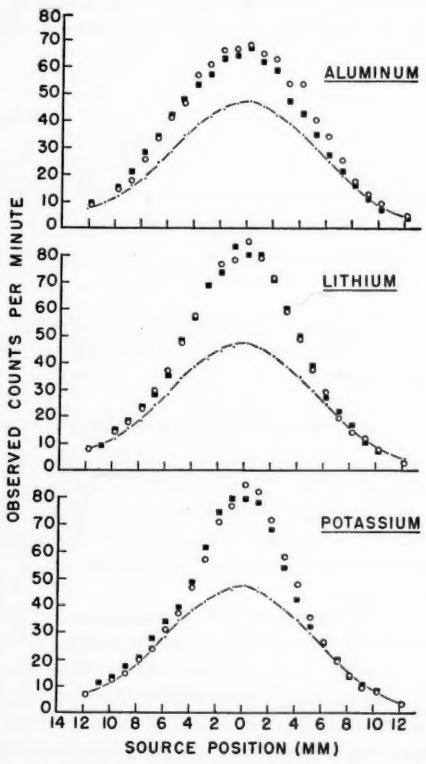


FIG. 2. Observed total angular correlation of annihilation photons from Al, Li, and K specimens clamped to the positron emitter. Separate runs are indicated by square and circular points. The dashed curve is the observed counting rate with no specimen metal against the emitter.

[5] Designed by R. E. Bell and R. L. Graham of Atomic Energy of Canada Ltd., Chalk River, Ontario.

FIG. 3. Emitter and specimen arrangement for shielded emitter method.

The electronic circuits were operated from Thorensen regulators. In addition, the channel counting rates were maintained constant by manually adjusting the lower discriminator settings every fifteen minutes. This small regular readjustment corrected for gain variation in the photomultipliers and amplifiers and for the different gamma-ray attenuation in the various specimens. Because the channel counting rate varied rapidly with the lower discriminator setting, the average energy of the accepted pulses changed very little over the range of this adjustment.

The usual experimental procedure was to move the source vertically in 1-mm steps and to count for 20 minutes at each position. The maximum coincidence rate was about 100 cpm. Runs were made with various specimen metals in position and also with no specimens. These blank runs were used to determine the fraction of positrons stopping in the emitter.

In Fig. 2 are shown the observed angular correlations for some typical runs with potassium, lithium, and aluminum. The corresponding blank run is shown by the dashed line under each curve.

A second method of arranging the source was suggested by Dr. E. W. Guptill. In this method the annihilation radiation from the emitter is shielded from the detectors by blocks of lead (Fig. 3). Positrons from the emitter traverse about ½ in. of air to reach the specimen which can be seen by the detectors. The emitter, shield, and specimen were raised and lowered as a unit.

Although this arrangement reduced the counting rate considerably, it had the advantage that the background was reduced by a much larger factor. In Fig. 4 are shown the results for calcium and beryllium from this type of experiment. The dashed line indicates the background, caused largely by scattered positrons.

III. THEORETICAL CALCULATION OF ANGULAR CORRELATION CURVES

To conserve momentum the angle between the two gamma rays from electron-positron annihilation must in general depart from 180° if the center of mass of the two particles was in motion. For low velocity the departure from 180° is $\phi = 2v/c$, where v is the transverse velocity of the center of mass of the particles and c is the velocity of light. It has been shown[6] that most positrons entering metals are thermalized in a time which is short compared with measured lifetimes. At thermal energies (1/40 ev) the positron velocity is very

[6] G. E. Lee-Whiting, Phys. Rev. **97**, 1557 (1955), and R. L. Garwin, Phys. Rev. **91**, 1571 (1953).

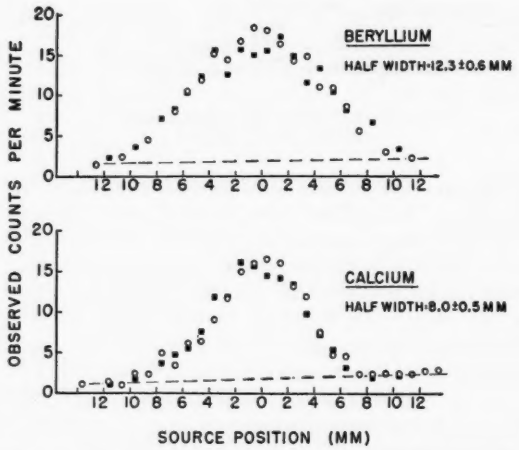


FIG. 4. Angular correlation of annihilation photons from Be and Ca observed by the shielded source method. The dashed curve is the background counting rate. Results of separate runs are indicated by round and square points.

small compared with conduction electron velocities, hence we can assume that the positron is at rest and that any motion of the center of mass of the annihilating pair is due to the velocity of the electron. Furthermore, DeBenedetti *et al.*[2] have pointed out that low-energy positrons would be expected to annihilate predominately with the conduction electrons: the Coulomb repulsion prevents the positrons from appreciably interacting with the inner electrons in the various elements.

In what follows we calculate the angular correlation which we expect to observe for positrons annihilating with the electrons in various metals assuming that the positrons are stationary and that the electrons have the momentum distribution predicted by the free electron theory.

1. Geometrical Resolution

The first step is to calculate the resolution function $R(z)$ due to finite thickness of source and detectors. (In the following analysis, the origin of the Cartesian coordinate system is at the mid-point of a line joining the two detectors; see Fig. 5. This line is the x-axis. The z-axis is vertical.) For a point source the resolution function, $R_0(z)$, is an isosceles triangle (Fig. 6) the width of the base of which is equal to the detector crystal thickness. The extension to a source of finite thickness, t, is merely the summation of $R_0(z)$ for

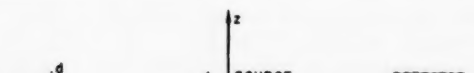


FIG. 5. Coordinate system with respect to source and detectors. For the arrangement used $d=7$ mm, $t=3$ mm, $s=2.45$ m.

various displacements of the point source. To simulate the actual source we assume that the activity is deposited uniformly over a ½-in. disk and that the detectors see this disk at about 10° (see Sec. II). For these conditions, a plot of source strength, $I(z)$, against vertical distance is an ellipse as shown in Fig. 7. The resolution function may then be found by integrating $R_0(a)$ times $I(a-z)$ over the height of the source, i.e.,

$$R(z)=\int_{-t/2}^{+t/2} R_0(a)I(a-z)da.$$

$R(z)$ is shown in Fig. 6. The resolution function was also determined experimentally by moving the detectors so close to the source that the width caused by angular divergence of the gamma rays was easily allowed for by a small correction.

The detector and source widths introduce a horizontal resolution function similar to $R(z)$. However, an additional effect must be included. The measured efficiency of detection at the edges of the 2-in. crystals is about 0.7 of the efficiency in the center. The geometric resolution function corrected for counter efficiency gives an effective horizontal resolution function,

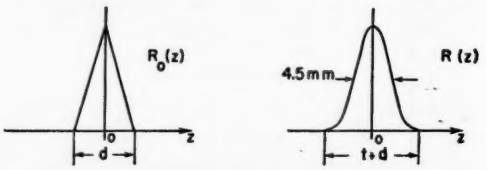


FIG. 6. Resolution functions $R_0(z)$ and $R(z)$ for a point source and a source of finite thickness.

$Y(y)$. As compared to a half-width (width at half-maximum) of 4.5 mm for $R(z)$, this function has a half-width of 29 mm.

2. Effect of Electron Motion

To calculate the effect of electron velocities we first recall that in the free electron theory of metals the tips of the momentum vectors of the conduction electrons uniformly fill a sphere. The volume of the sphere is proportional to the density of the free electrons in the metallic crystal.

If we consider annihilation radiation arising from electrons which have only one value, v_z, for the z component of velocity and v_y for the y component of velocity, we would observe experimentally a correlation having the shape of the $R(z)$ function times the $Y(y)$ function but shifted vertically from the origin an amount $sv_z/2c$ and horizontally an amount $sv_y/2c$, where s is the distance from source to detectors. In a free electron gas the number of electrons having a y and z component of velocity between v_y and v_y+dv_y and v_z and v_z+dv_z is given by

$$N(v_y,v_z)dv_ydv_z \propto (v_F^2-v_y^2-v_z^2)^{\frac{1}{2}}dv_ydv_z,$$

where v_F is the maximum velocity corresponding to the Fermi energy. To obtain the predicted curve for the conduction electrons we fold the RY product function with the electron distribution function, $N(v_y,v_z)$. The integrated function,

$$T(z,y)=\int_{-v_F}^{+v_F}\int_{-(v_F^2-v_z^2)^{\frac{1}{2}}}^{+(v_F^2-v_z^2)^{\frac{1}{2}}} N(v_z,v_y) \times R(z-sv_z/2c)Y(y-sv_y/2c)dv_ydv_z,$$

is proportional to rate of coincidences expected between the two detectors of Fig. 5 as the source of annihilation radiation is moved in the $y-z$ plane. Since we are interested in this curve for $y=0$ only we shall write simply $T(z)$. Curves calculated in this manner for various maximum electron velocities, corresponding to various metals, are shown in Fig. 8 and the half-widths are tabulated in Table I.

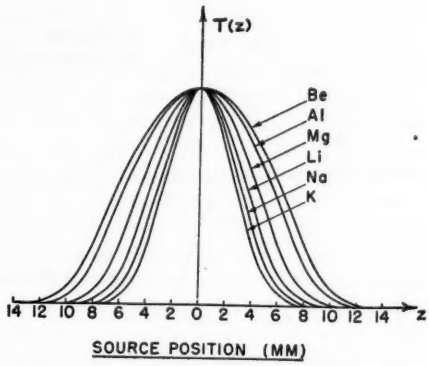


FIG. 8. Calculated angular correlation curves for various metals

IV. ANALYSIS OF EXPERIMENTAL RESULTS

An observed angular correlation curve is the sum of curves representing annihilations which took place in the emitter and annihilations which occurred in the specimen. We assume that the area of the blank curve is proportional to the number of positrons that did not escape from the emitter and that the area of the curve observed with a specimen in position is proportional to the total number of positrons. Hence by subtraction of areas we can find the fraction of positrons which entered the specimen metal. The fraction which annihilated in the specimen is somewhat less than this because some of the positrons were scattered back to the emitter.

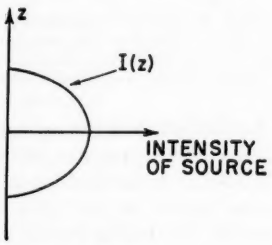


FIG. 7. Variation of source strength through the thickness of the source.

It was observed that the areas obtained from the many blank runs varied by about ±3 percent and that the areas obtained with the various specimens fluctuated about ±6 percent. As no definite cause could be assigned to this variation, which was a few times the statistical uncertainty, it was decided to take the average area of the blank curves to be proportional to the annihilation rate of positrons which do not leave the emitter. Similarly, the average area of the various specimen curves was taken to be proportional to the total annihilation rate. The specimen curves were then normalized to their average area before subtracting the curve representing annihilations in the emitter. For example, the procedure for an aluminum run was first to normalize by increasing the area by 100/92 (Table II). To determine the amount to be subtracted, we see from the table that about 64 percent of the annihilations occur in the emitter so that only about 36 percent of the positrons enter the specimen. Seliger[7] showed that a considerable fraction is reflected back to the emitter. After summing for repeated reflections, the net fraction returned to the emitter is found to be 0.22. Thus the fraction of positrons annihilating in the emitter when covered with an aluminum specimen is 64 percent plus 0.22×36 percent=72 percent=1.12×64 percent. Hence from the normalized aluminum curve we subtract 1.12 times the measured blank curve.

Typical net curves are shown in Fig. 9. These curves may be compared with similar curves shown in Fig. 4 which were obtained by the shielded emitter method. The observed and calculated half-widths are listed in Table I.

In Fig. 10 is plotted the Fermi energy of the electrons of the specimen metal against the width of the correlation curve. The full line is the calculated curve and the points are observed values. The specimen was not tilted when used with the shielded emitter and hence

TABLE I. Calculated and observed half-widths of angular correlation curves for various metals. Resolution function, $R(z)$, half-width=4.5 mm.

	Free electrons per atom	Calculated[a] Fermi energy (ev)	Calculated half-width (mm)	Observed half-widths (mm) Subtracting blank method	Observed half-widths (mm) Shielded emitter method	Fermi energy from measured half-widths (ev)
K	1	2.04	6.0	6.0±0.2		2.1±0.3
Na	1	3.16	6.7	6.8±0.3		3.3±0.4
Li	1	4.75	7.8	7.7±0.3		4.6±0.5
Ca	2	4.72	7.8	7.7±0.3	8.0±0.5	4.8±0.4
Mg	2	7.16	9.3		9.4±0.3	7.4±0.5
Al	3	11.75	11.8	11.7±0.6	11.1±0.5	11.1±0.9
Be	2	14.3	12.9	12.8±0.6	12.3±0.6	13.6±0.9

[a] These Fermi energies have been calculated by assuming that the free electrons have the entire volume of the metal in which to move. If the ion core volume is excluded, the resultant Fermi energy will be higher by several percent for the heaviest of these metals.

[7] H. H. Seliger, Phys. Rev. 88, 408 (1952).

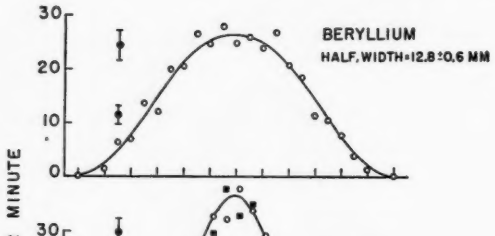


FIG. 9. Typical net curves resulting from subtraction of blank curve from the total coincident counting rate curves. The full line is the calculated curve normalized to equal area. The statistical uncertainty of the individual points is indicated by the vertical bars at the side of the curves.

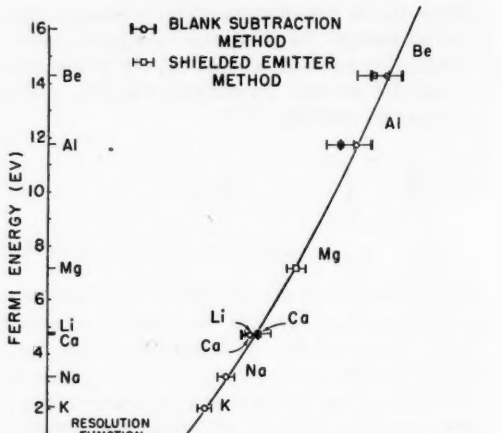


FIG. 10. Plot of Fermi energy against calculated half-width of correlation function. The measured half-widths are shown on the curve, the square and round points representing results of the shielded emitter and blank subtraction methods, respectively.

the curves calculated for this method are slightly narrower. However, the difference is within the experimental error, so that for clarity the calculated curve for this method has been omitted from the figure.

V. DISCUSSION

1. Experimental Uncertainties

On the assumption that the fluctuations in area have been due primarily to changes in coincidence efficiency, we have normalized all observed curves to the same area. Fortunately the widths of the net curves are not strongly dependent upon the method of normalizing. It can be shown for the narrow curves (K, Li) that the maximum change, 11 percent, in the area of the observed curve will cause only about a 4 percent change in the half-width of the net curve. The experimental error is also about 4 percent, and hence the mean uncertainty from the normalization is about half the statistical error shown in the table. For the wider curves the uncertainty will be less because both the total and blank curves have approximately the same width.

TABLE II. Typical areas of observed curves and effective positron reflection coefficients.

	Observed areas (relative values)	Net positron reflection coeff.	Percentage of positrons stopping in emitter
Observed area for Al	92	0.22	$64+36\times0.22=64\times1.12$
Observed area for Ca	110	0.26	$64+36\times0.26=64\times1.14$
Observed area for K	95	0.26	$64+36\times0.26=64\times1.14$
Average area of all runs	100		
Area of blank runs	64		
Percentage of positrons entering specimen	36		

The back-scattering was measured by Seliger[7] for positrons from a thin emitter whereas the positrons being reflected from the specimen in the present experiment have been degraded in energy by passing through the aluminum foil of the emitter. However, the error caused by this uncertainty cannot be very large because even a 50 percent change in the reflection coefficient will not change the half-width by more than the experimental error.

In the shielded emitter method, the background to be subtracted is in part caused by back-scattered positrons. Hence it is different for the various specimens and cannot be accurately determined merely by removing the specimen. The error assigned to these results contains an estimate of the background uncertainty.

Small-angle scattering of gamma rays (Rayleigh scattering) has been estimated and found to contribute negligibly to the broadening of the observed coincidence function.

2. Theoretical Assumptions and Implications

It is immediately evident from the experimental results that the two photons from electron positron

annihilation in light metals carry away an amount of momentum which is of the same order as the momentum of the conduction electrons. Since it has been shown that positrons thermalize very quickly,[6] we are led to believe that they annihilate chiefly with the conduction electrons of these metals.

If annihilations are assumed to occur with conduction electrons only, and to occur by a process which makes the total momentum of the two photons equal to the momentum of the annihilating electron, the quantitative experimental results may be interpreted in two ways: *Firstly*, we may assume that the momentum distribution of the conduction electrons is described by the free electron theory and deduce the velocity dependence of the annihilation cross section. The full line of Fig. 10 is calculated on this basis for a $1/v$ cross section. This corresponds to the assumption of equal probability of annihilation with all conduction electrons which was implicit in the arguments in Sec. III. If it were assumed that the annihilation cross section varies as $1/v^2$, the calculated half-widths would be too small by about three times the experimental error. Similarly, an annihilation cross section which was assumed to be independent of velocity would lead to calculated half-widths which are correspondingly large. *Secondly*, we may assume that the cross section varies as $1/v$ and use the experimental results to determine the Fermi energy of the conduction electrons. The Fermi energies so determined are listed in Table I. The results agree well with the values calculated from the free electron theory. The assumption of an annihilation cross section which varies other than as $1/v$ would yield Fermi energies which differ considerably from calculated values.

The mechanism of annihilation is not well understood. The most obvious process is the ordinary annihilation of an electron and positron in a collision, for which the cross section was first derived by Dirac.[8] However, as Bell and Graham have pointed out,[9] the cross section for direct annihilation leads to a lifetime several times that observed. In fact, for light metals, the electron density is not sufficient to account for the observed lifetime even if all the electrons are available for annihilations. It has been suggested[10] that Coulomb attraction may cause an increased electron density near the positron and hence a shorter lifetime. If this effect can be shown to shorten the calculated lifetime sufficiently, then the lifetime and angular correlation measurements will be consistent. If not, it may be necessary to consider the formation of positronium in metals in a time short compared with $\sim 10^{-10}$ sec, the observed lifetime. The atom might then annihilate in a time of the order of the singlet lifetime, 1.25×10^{-10} sec. This picture is not free from obstacles. It would be necessary to have a mode of formation in which the positronium atom retains nearly all of the conduction electron momentum and in which therefore, a photon or possibly an Auger electron removes the excess energy. In addition, the atom must retain its kinetic energy during collisions in the metal. Other pictures of varying complexity are also possible. The unsolved problem remains: to account for the observation that approximately 10^{-10} sec after a fast positron enters a metal, annihilation photons emerge with a total momentum which appears to be that of a conduction electron.

ACKNOWLEDGMENTS

The determination of electron velocities by measuring the width of the angular distribution curve instead of the "tail," as had been previously attempted, was suggested to us by Dr. G. E. Lee-Whiting to whom we are indebted for many helpful discussions. It is also a pleasure to acknowledge the advice and assistance of Dr. R. L. Graham and Mr. R. M. MacIntyre who contributed to the preliminary experiments. Many others of the Physics Department of Dalhousie University and the Physics Division of Atomic Energy of Canada Limited have assisted through discussions.

We wish to thank the Atomic Energy of Canada Limited for loan of equipment and the National Research Council for financial support. This work was done while one of us (R. E. Green) was the holder of a MacGregor Teaching Fellowship at Dalhousie University.

[8] P. A. M. Dirac, Proc. Cambridge Phil. Soc. **26**, 361 (1930).
[9] R. E. Bell and R. L. Graham, Phys. Rev. **90**, 644 (1953).
[10] Reference 2, footnote 18.

Magnetic Scattering of Slow Neutrons by Gaseous Oxygen*

H. PALEVSKY AND R. M. EISBERG
Brookhaven National Laboratory, Upton, New York

(Received January 5, 1954)

The scattering of neutrons by oxygen was measured as a function of wavelength from 2.5–10 A. By subtracting the calculated nuclear contribution the magnetic scattering is established. The magnitude of the magnetic interaction at long wavelengths is found to be in agreement with the theory of Halpern and Johnson. The measurement is not very sensitive to the details of the spatial distribution of the magnetic electrons. However, the observed wavelength dependence of the magnetic scattering is in agreement with the calculations of Kleiner based on the oxygen wave function derived by Meckler.

I. INTRODUCTION

THE scattering of slow neutrons by gaseous oxygen arises from both nuclear and magnetic interactions. The latter is of particular interest because oxygen is the simplest molecular system exhibiting paramagnetism, and there is a strong possibility that the magnetic scattering can be predicted from fundamental theory. In the past, magnetic measurements (as susceptibility) have dealt with the bulk or macroscopic properties of this gas. High-intensity neutron beams, however, now make it possible to perform experiments that probe the magnetic properties on a molecular (microscopic) scale. Although the scattering of slow neutrons by magnetic substances in general depends in a complicated way on intermolecular forces, in oxygen gas the intermolecular effects are extremely small and the magnetic interaction is between the neutron and the "free" magnetic moment of the oxygen molecule. This latter interaction has been theoretically formulated by Halpern and Johnson.[1]

In order to determine the magnetic scattering cross section it is necessary to subtract the nuclear effects from the total measured cross section. Fortunately the nuclear properties of oxygen are well suited for this purpose. The neutron absorption cross section is very small and the spin and isotope incoherence are negligible (zero spin and almost monoisotopic). For these reasons the coherent nuclear scattering amplitude, which is needed for calculating the nuclear effects, may be determined by a free-atom cross-section measurement at neutron energies of several electron volts,[2] where magnetic scattering is negligible.

II. METHOD

The present experiment is performed by measuring the transmission of neutrons through gaseous oxygen, which gives the total neutron cross section. Because the magnetic moments of the oxygen molecules are randomly oriented, there is no coherence between nuclear and magnetic scattering and the total cross section is separable into magnetic and nuclear terms. The contribution of the nuclear terms has been accurately calculated by Halpern and Appleton.[3] An estimate of the ratio of the magnetic to nuclear cross sections can be arrived at by simple considerations. At wavelengths large (~10 A) compared to the interatomic spacing of the oxygen atoms, the form factor for the nuclear and magnetic scattering approaches unity. For oxygen molecules at rest, in the limit of long wavelength, the nuclear scattering cross section of the molecule is approximately four times σ_b, the nuclear cross section of a bound oxygen atom.

The cross section for the magnetic interaction is given by[3]

$$\sigma_m=\frac{8\pi}{3}\left(\frac{\mu}{m}\right)^2 S(S+1)\left[\frac{\gamma^2e^2}{mc^2}\right]|F|^2, \tag{1}$$

where m is the mass of the neutron, μ is the reduced mass of the oxygen molecule-neutron system; S the electronic spin of the oxygen molecule; γ the magnetic moment of the neutron, expressed in Bohr magnetons and e^2/mc^2 the classical electron radius.

$$|F|^2\equiv\frac{1}{4\pi}\int|f(k)|^2\sin\theta d\theta d\varphi$$

is the integral form factor, normalized to approach unity as the neutron momentum $k\to0$. Thus, in the long wavelength limit, the magnetic cross section is given by

$$\sigma_m{}^0=\frac{8\pi}{3}\left(\frac{\mu}{m}\right)^2 S(S+1)\left[\frac{\gamma^2e^2}{mc^2}\right]^2, \tag{2}$$

and the ratio

$$\frac{\sigma_{\text{magnetic}}}{\sigma_{\text{nuclear}}}\xrightarrow{\quad}\frac{2\pi}{3\sigma_b}S(S+1)\left[\frac{\gamma^2e^2}{mc^2}\right]^2=\frac{1}{3.5}. \tag{3}$$

As the Doppler motion (resulting from the finite sample temperature) affects the magnetic and nuclear scattering in the same way, the ratio is not changed by the Doppler effect.

In the long wavelength limit, a one percent error in the total cross-section measurement would result in about a

* Work carried out under contract with the U. S. Atomic Energy Commission.

[1] O. Halpern and M. H. Johnson, Phys. Rev. **55**, 898 (1939).

[2] E. Melkonian, Phys. Rev. **76**, 1750 (1949).

[3] O. Halpern and G. L. Appleton, Phys. Rev. **90**, 869 (1953).

4.5 percent error in the magnetic cross section, because of their relative magnitude. As the wavelength is decreased, the magnetic scattering decreases more rapidly than the nuclear scattering until at wavelengths <1 A the magnetic scattering is less than 2 percent of the nuclear scattering. For this reason the calculations of Halpern and Appleton were made in regions between 5–10 A, where one may hope to obtain the magnetic scattering with good accuracy from a total cross-section measurement. For the sake of completeness the experimental data were taken over a wavelength range from 2.5–10 A. The data for wavelengths less than 5 A can be compared with existing measurements.

III. EXPERIMENTAL DETAILS

Figure 1 shows the experimental apparatus. A collimated beam of neutrons from the Brookhaven pile is chopped into a series of approximately 200-μsec bursts by the "slow chopper," a rotating cadmium shutter. The beam passes through a filter which consists of four inches of Be maintained at liquid N_2 temperature. The Be transmits, with little attenuation, neutrons of wavelengths longer than 3.9 A but is essentially opaque to neutrons of shorter wavelength. Thus the filter eliminates neutrons of thermal and higher energies from the beam and thereby greatly reduces background. Next the beam passes through the sample vessel which is 3.5 meters long. The vessel may either be evacuated or filled to 10 atmospheres with O_2 and the total cross section of the O_2 is obtained by measuring the attenuation of the beam upon filling the vessel. The neutrons are detected in the shielded BF_3 counter, which is located 5.0 meters from the chopper, in a good geometry arrangement with respect to the sample vessel. The velocities or wavelengths of the detected neutrons are determined by measuring the flight time over the 5.0-meter path.

The chopped beam is monitored by a BF_3 counter which is in a beam obtained from the same region of the pile as the chopped beam. The monitor beam is filtered through graphite so that the mean wavelength of the neutrons detected by the monitor lies within the wavelength range 5 to 10 A that is being investigated with the chopper.

The oxygen used to fill the sample vessel is obtained from the electrolysis of H_2O, the only impurity being ~1 percent hydrogen. The hydrogen is removed by passing the gas over heated CuO which catalyzes the oxidation of H_2 to form H_2O. The gas is than passed through a dry ice trap which removes the H_2O. The remaining impurities in the sample are traces of unoxidized H_2, and H_2O resulting from the vapor pressure at dry ice temperatures. Typical samples of gas from the target vessel were analyzed for H_2 and H_2O. These impurities were found to be present to the extent of 0.01 mole percent. Since the molecular cross sections of H_2 or H_2O are about ten times the molecular cross sections of O_2 at these wavelengths, the error in the measured O_2 cross section due to sample impurities is 0.2 percent.

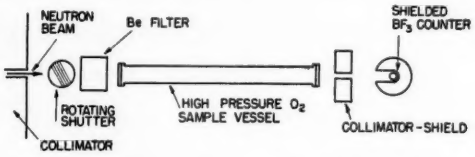


FIG. 1. Experimental arrangement of apparatus.

To calculate the cross section from the measured transmission the number of molecules/cm² in the sample is required. Measurements of the temperature and pressure of the gas and the length of the target vessel were made with $\frac{1}{4}$ percent accuracy. The molecular density was computed by means of the Beattie-Bridgeman equation of state. At the operating pressure of 10 atmospheres this equation of state indicates a 0.7 percent deviation from the perfect gas law.

The time-of-flight spectra of the neutrons in the beam is analyzed by a 12-channel delayed-coincidence analyzer. Flight times are measured relative to a one-megacycle crystal oscillator. The zero of time is set by a calibration procedure referring to the crystal lattice spacing of graphite[4] which is known from x-ray diffraction experiments. The neutron time of flight is measured with an accuracy of about $\frac{1}{4}$ percent.

The transmission is given by the ratio of the counting rate less background with O_2 in the beam divided by the counting rate less background of the open beam. Backgrounds are measured by inserting a 0.010-in. cadmium foil in the beam. This foil absorbs more than 99 percent of the timed neutrons but does not affect the epicadmium background neutrons. The data were obtained by repeatedly running the four measurements involved in the measurement of the transmission. The cross section was calculated from the measured transmission and molecular density of the sample. The results are shown in Fig. 2. The abscissa is neutron wavelength including a small correction for the finite resolution of the spectrometer and the ordinate is the total cross section per O_2 molecule. The indicated errors range from ± 1.6 percent at 10 A to ± 0.1 percent at 5 A. In order to compare the results of the present experiment with existing cross section information, which extends up to 5 A, the cross section was measured from 2.5 A to 5 A with somewhat poorer statistics. These data are in good agreement with the measurements of Melkonian.[2]

IV. CONCLUSIONS

Figure 3 shows the experimental magnetic scattering in the region from 5–10 A, obtained by subtracting the calculated nuclear scattering[3] from the measured points. The solid line is a theoretical prediction of the magnetic

[4] Carter, Palevsky, Myers, and Hughes, Phys. Rev. **92**, 716 (1953).

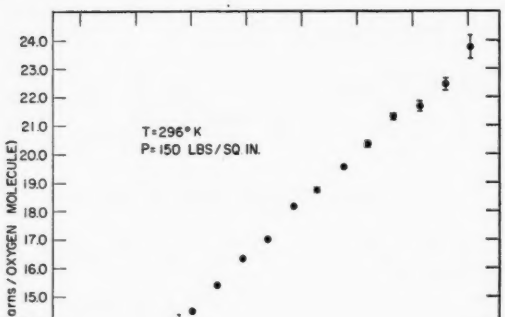


FIG. 2. Total cross section of oxygen in the wavelength region 2.5–10 A. Where no error is shown, the statistical error is equal to or less than the size of the plotted point.

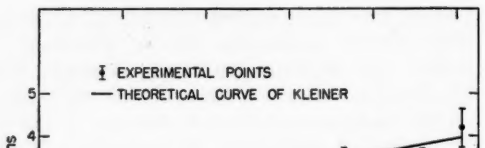


FIG. 3. Magnetic scattering of oxygen in the wavelength region 5–10 A. Where no error is shown, the statistical error is equal to or less than the size of the plotted point.

scattering made by Kleiner[5] using the oxygen wave function of Meckler.[6] In the region 7.5 to 10 A, Kleiner's calculations indicate that for any reasonable oxygen wave functions the magnetic form factor is very nearly unity. In this wavelength interval the slope of the magnetic scattering depends only on the limiting cross section σ_m^0 [Eq. (2)] and the temperature of the gas. Evaluation of Eq. (2) gives $\sigma_m^0=4.58$ barns, and use of this value in calculating the Doppler effect gives a slope of 0.356 barn/A.

In order to compare the experimental results with those theoretically expected, a least squares fit to the six experimental points was made between 7.7 and 10.1 A gives an experimental slope of 0.340 ± 0.007 barn/A, which is about 4.5 percent lower than the theoretical prediction. This discrepancy is to be compared to the 2 percent standard deviation in the least-squares fit of the experimental data.

The magnetic scattering between 5 and 7.5 A depends not only on the σ_m^0 and temperature but also to a small extent on the details of the spatial distribution of the magnetic electrons in the molecule, i.e., in this wavelength region there is a departure of the form factor from unity. The experimental points here are known with about one percent (standard deviation) accuracy; however, they all lie above the theoretical curve of Kleiner. The average deviation is seen (Fig. 3) to be about 6 percent. Even though the form factor in this wavelength region is not unity, the work of Kleiner indicates that the approximate wave functions he used should accurately predict the magnetic scattering. Both the discrepancy in the slope and magnitude of the experimental cross sections as compared to theory could be explained by a small (~1 percent) error in the calculation of the nuclear scattering.

This experiment establishes the magnetic scattering of neutrons by oxygen gas. The slope of the magnetic cross section curve verifies that at long wavelengths σ_m^0 approaches the value predicted by theory for the interaction of the neutron and a "free" magnetic moment. The measurement is rather insensitive to the details of the spatial distribution of the magnetic electrons because a transmission experiment depends on the integral of the form factor over all angles. The details are further smeared out by the Doppler motion of the gas. In the region between 5 and 10 A, these two effects make it very difficult to obtain with any accuracy properties of the distribution of the magnetic electrons in oxygen. The smearing arising from the Doppler motion becomes less serious at shorter wavelengths and it would be possible to learn more about the magnetic moment distribution in the molecule from data between 2 and 5 A if the nuclear contributions could be calculated.

The authors are indebted to Martin E. Rickey for his able assistance in carrying out the measurements. We also wish to thank Professor T. I. Taylor for suggesting the method for purifying the oxygen and performing the analysis for impurities. The discussions with Dr. W. Kleiner, Dr. L. Van Hove, and Dr. D. J. Hughes are gratefully acknowledged. We are especially indebted to Dr. Hughes for his aid in the early planning of the experiment and his assistance in the preparation of the final manuscript.

[5] W. Kleiner, Phys. Rev. **97**, 411 (1955).
[6] A. Meckler, J. Chem. Phys. **21**, 1750 (1953).

Electron Impact Broadening of Spectral Lines*

B. KIVEL,† S. BLOOM,‡ AND H. MARGENAU
Sloane Physics Laboratory,§ Yale University, New Haven, Connecticut
(Received August 13, 1954)

Ion densities in electrical discharges and in stars are often determined by observing the widths of spectral lines emitted or absorbed. The Holtsmark theory, usually employed in the interpretation of the data, restricts itself to ions and ignores electron collisions, which are treated in a qualitative way. This paper describes a detailed theory of electron effects upon the Lyman alpha line.

A convenient physical classification of the result is as follows. Electrons produce, first of all, what we have called a *universal* line effect, a broadening occurring regardless of the nature of the atomic transition. Its cause is the transfer of energy from the electrons to one or the other of the two atomic states via scattering. So far as the analogy is proper, this effect is what is sometimes called "second order Stark broadening." It forms the predominating feature when a linear Stark effect is absent.

Secondly, electrons can enforce the optical transition by collisions of the second kind. This mechanism, called *quenching*, broadens the line. It alone has the property of reducing the intensity of the line under normal conditions.

A third possible contribution to the line width results from *polarization*, the counterpart of a first-order Stark effect. It comes about if one of the atomic states (the upper in our example) is degenerate or nearly degenerate. The electron, while flying past the atom, can then induce transitions without much energy transfer, i.e., polarize the atom. A distinction is made between (a) polarization through induction of a temporary dipole moment and (b) polarization through reorientation of an atomic dipole already present by virtue of a neighboring ion. The first of these is a large effect; its calculation requires the removal of divergences which are ordinarily encountered in the matrix elements.

The summary at the end of the paper contains numerical results on the broadening of L_α by these several agencies under conditions approximating the solar photosphere. They are all small in comparison with Holtsmark broadening by ions. But the example is unusual; in general the electron effects are far from negligible.

I. INTRODUCTION

SPECTRAL lines emitted in a plasma are known to be broadened in a manner characteristic of the ion density and the temperature in the discharge. Indeed, line widths are often regarded as "probes" by means of which these two parameters can be determined. Tests of this kind are based upon the theory of line broadening developed long ago by Holtsmark,[1] a theory which assumes that the ions exert Stark effects whose magnitudes vary as the ions move, thereby diffusing the line intensity. The electrons, which are as numerous as the ions, are supposed to produce small effects because of their greater velocity and their ubiquity: they have less time to interact with the radiating atoms, and by being nearly everywhere they affect the atoms like a static distribution of charge, which shifts the lines but does not broaden them.

Despite the plausibility of these considerations, the Holtsmark theory is plagued by an insecurity arising from its failure to include the electrons in its concerns. This has long been felt by astronomers and physicists alike, and many qualitative attempts have been made to treat the electrons. Some of these have had wide attention and interesting use, for the rigorous theory is cumbersome, and the likelihood that it will give answers wholly different from those intuitively expected is small. Probably for these reasons, such a theory has not previously been worked out. In the present paper we undertake to do this, not in the hope of spectacularly new results, but to provide a carefully analyzed instance against which simpler models can be tested. Our approach is via the theory of natural line width; it involves old and familiar ideas. We are not sure, however, that it is the most elegant and the most direct attack upon the problem; in fact it seems to us at times as if our results should be obtainable in simpler ways.

The usual theories of pressure broadening are not applicable to perturbing electrons because the uncertainty principle prevents electrons from pursuing a classical path along which an interaction potential can be defined. It is necessary, therefore, to incorporate them into the quantum mechanical system from the very beginning, a process which leads to the elaborate treatment developed in this paper. Some of the results obtained are intuitively plausible and almost derivable from various simple, classical models. But the consistency of the guesses based on such considerations is difficult to establish unless a comprehensive account is at hand. Aside from being systematic, our treatment also indicates many connections with other physical phenomena, such as scattering, quenching of excited states, etc., and thereby deepens understanding.

To keep the work simple and the methods clear, no extensive applications are given in the present paper. The magnitudes of the various effects are illustrated in the terms of one simple example, the width of the first Lyman line in atomic hydrogen under typical discharge conditions. The analysis dealing with "universal broadening" (by which we mean a transfer of energy to the atomic state via electron scattering in a manner computable without assuming a change in atomic states; this mechanism is always present, whether there is degeneracy of atomic states or not) in Sec. III is limited

* Part of a dissertation presented for the degree of Doctor of Philosophy in Yale University.
† Now at Los Alamos Scientific Laboratory, Los Alamos, New Mexico.
‡ Now at Radio Corporation of America, Princeton, New Jersey.
§ Supported by the Office of Naval Research and the U. S. Atomic Energy Commission.
[1] A. B. Underhill, Astrophys. J. **116**, 446 (1953).

to line widths comparable to the natural width. Since this effect is found to be small and our applications fall within the limits of validity of our treatment, we have not thought it necessary to extend it. Subsequent publications will deal with more practical applications, among them the contribution of plasma electrons to the width of the Balmer lines.

Current investigations of hydrogen line breadths in stellar spectra indicate that Stark broadening by ions is insufficient to account for the observations.[1,2] They suggest that the diffculties may be obviated by taking into account the electron effects. Another discrepancy is the failure of Holtsmark theory for ions to predict the large shifts of argon spectral lines found by Kantrowitz[3] in the study of shock phenomena. Baranger[4] attributes the difference to electrons. The electron effect also interests workers using spectral line widths to measure plasma characteristics in discharge tubes.[5] To account for it the density of charge employed in the Holtsmark theory is often taken to be intermediate between the ion density and ion plus electron density.[6]

Reasonable qualitative conclusions about the electron broadening and shift of spectral lines can be drawn from a semiclassical model.[7] If, for example, the electron density is low and the lifetime of the excited radiator is short, then not many of the atoms will be hit before they emit, if by a hit we mean a "sufficiently close" collision. In the present paper we find the cross section of this collision; its radius turns out to be essentially the de Broglie wavelength of the electron. Thus high-energy electrons, which produce small effects, may be considered as passing the atom too quickly to polarize it or to be scattered by it.

For the quantitative treatment of the problem quantum theory of radiation is used. On the basis of the results obtained it is possible, first of all, to delimit the large region of high temperatures and low densities where electron effects are minor and ignorable. With decreasing electron temperature and increasing electron density, first-order corrections to the natural line become important and are found. Here two effective broadening mechanisms are distinguished: (1) change in the electron energy during a radiative atomic transition (*universal broadening*) and (2) distortion of the atomic state-function (*polarization*) by the electrons. This second effect is important only if one of the atomic levels is strongly coupled to a neighboring level. Application is made to the hydrogen-like $2p$ to $1s$ transition. By inclusion of the $2s$ state, the effects of polarization of the excited state by electrons can be studied. Attention will also be given to polarization in the physically interesting situation of a hydrogen atom amid numerous positive ions.

II. GENERAL CONSIDERATIONS[8]

Electron broadening can be studied as an extension of the Hoyt and Weisskopf-Wigner treatment for the natural line width.[9] The natural decay for a two-state atom ($E_1>E_0$) is described by the growth equations,

$$i\hbar\dot{d}=\sum_r J_r^* b_r e^{i\omega t}, \tag{1}$$

and

$$i\hbar\dot{b}_r=J_r d e^{-i\omega t}, \tag{2}$$

where $\omega=\omega_{\text{atom}}-\omega_r$, $\hbar\omega_{\text{atom}}=E_1-E_0$, and $\hbar\omega_r$ is the energy of the emitted radiation. In these equations the matrix elements for the atom-radiation interaction (J_r) couple the probability amplitudes for the state: excited atom and dark field (d) with: atomic ground state and a single photon (b_r). For exponential damping ($d=e^{-\gamma t}$), Eq. (2) yields the natural intensity distribution (ρ_r is a weighting function)

$$I(\omega_r)=\rho_r|b_r(t=\infty)|^2=\rho_r|J_r/\hbar|^2/(\gamma^2+\omega^2). \tag{3}$$

To include a perturber ($\mathbf{R}$) with the atom and the radiation field, we write the Schrödinger equation

$$i\hbar(\partial/\partial t)\Psi(\mathbf{r},\mathbf{R})=H(\mathbf{r},\mathbf{R})\Psi(\mathbf{r},\mathbf{R}). \tag{4}$$

The Hamiltonian (H) includes terms for unperturbed atom and radiation field $[H_a(\mathbf{r})]$, the perturber $[H_e(\mathbf{R})]$, an interaction between atom and perturber $[C(\mathbf{r},\mathbf{R})]$, and finally the interaction between atom and radiation $[J(\mathbf{r})]$; it neglects the interaction between perturber and radiation.

In our notation symbols referring to the radiation field are suppressed. The wave functions and eigenenergies for the field are understood to be contained in the atomic ψ_n and E_n.

We are here interested in the extreme of an undistorted atom; hence, we seek a solution which is an expansion of product functions with unperturbed atomic functions, i.e.,

$$\Psi=\sum_n\sum_\lambda{}^n a_{n\lambda}(t)\psi_n(\mathbf{r})\phi_\lambda{}^n(\mathbf{R})\exp[-i(\epsilon_\lambda{}^n+E_n)t/\hbar], \tag{5}$$

where $\sum_\lambda{}^n$ means the sum and/or integral over all perturber states λ while the atom is in state n. The atomic function satisfies

$$H_a(\mathbf{r})\psi_n(\mathbf{r})=E_n\psi_n(\mathbf{r}), \quad \int|\psi_n|^2 d\mathbf{r}=1; \tag{6}$$

the normalized electron function for an electron in the

[1] G. J. Odgers, Astrophys. J. **116**, 444 (1953).
[2] A. Kantrowitz, Phys. Rev. **90**, 368 (1953).
[3] M. Baranger, Phys. Rev. **91**, 436 (1953). See also P. Shultz, Ann. Physik **3**, 280 (1948).
[4] See H. N. Olsen and W. S. Huxford, Phys. Rev. **87**, 922 (1952).
[5] G. Jurgens, Z. Physik **134**, 21 (1952).
[6] See D. R. Inglis and E. Teller, Astrophys. J. **90**, 439 (1939).
[7] H. Margenau and S. Bloom, Oak Ridge National Laboratory Progress Report, March 1, 1952 (unpublished).
[8] See S. Rosseland, *Theoretical Astrophysics* (Clarendon Press, Oxford, 1936), especially Chap. VIII.

field of an atom in state ψ_n obeys

$$[H_e(\mathbf{R})+C_{n,n}(\mathbf{R})]\phi_\lambda^n(\mathbf{R})=\epsilon_\lambda^n\phi_\lambda^n(\mathbf{R}), \quad (7)$$

$$\int|\phi_\lambda^n|^2d\mathbf{R}=1, \quad C_{m,n}(\mathbf{R})\equiv\int\psi_m^*C\psi_n d\mathbf{r}.$$

To justify the use of the expansion (5), we note that the best product function with neglect of the exchange property is given by the Hartree equations:

$$\left(H_a+\int\phi_\lambda^{n*}C\phi_\lambda^n d\mathbf{R}\right)\psi_n^\lambda=E_n^\lambda\psi_n^\lambda; \quad (8)$$

$$\left(H_e+\int\psi_n^{\lambda*}C\psi_n^\lambda d\mathbf{r}\right)\phi_\lambda^n=\epsilon_\lambda^n\phi_\lambda^n. \quad (9)$$

For high perturber energies (ϵ_λ^n), ϕ_λ^n is to a good approximation a plane wave. Then $|\phi_\lambda^n|^2$ is a constant independent of λ, and ψ_n^λ is the atomic state in a uniform charge distribution normalized to the perturber density. At low perturber density it is proper to neglect the shielding effects of this charge distribution on ψ_n^λ. Thus we ignore the C matrix element in (8) and our equations take the form (6) and (7). Since the product $\psi_n\phi_\lambda^n$ is not a stationary solution for the material system, the coefficients will, of course, depend on C.

On substituting (5) into (4) and using (6) and (7), we find that

$$i\hbar\dot{a}_{m\mu}\exp[-i(\epsilon_\mu^m+E_m)t/\hbar]$$

$$=\sum_n\sum_\lambda{}^n a_{n\lambda}J_{mn}\int\phi_\mu^{m*}\phi_\lambda^n d\mathbf{R}\exp[-i(\epsilon_\lambda^n+E_n)t/\hbar]$$

$$+\sum_{(n\neq m)}\sum_\lambda{}^n a_{n\lambda}\int\phi_\mu^{m*}C_{mn}\phi_\lambda^n d\mathbf{R}$$

$$\times\exp[-i(\epsilon_\lambda^n+E_n)t/\hbar]. \quad (10)$$

This equation differs in two respects from the Weisskopf-Wigner equations. First it provides for an exchange of energy between electron and radiation field by the presence of the ϵ_μ^m and ϵ_λ^n factors in the exponent of the first expression on the right. These remain because ϕ_μ^m and ϕ_λ^n are not orthogonal. The likelihood of these exchanges or shifts depends on the integral factor,

$$\int\phi_\mu^{m*}\phi_\lambda^n d\mathbf{R}.$$

The second effect additional to the natural line width is inherent in the terms contributed by the C elements on the right of (10). They represent polarization or distortion of the atom when the states m and n are degenerate or nearly degenerate; if n and m denote the levels between which the radiative transition occurs these added terms describe quenching.

Henceforth, we speak of the first additional effect as universal broadening, of the second as polarization or quenching. The two may be treated separately in our case of low electron density and high temperature, where modification of the spectral line is small.

The first correction, to which the remainder of this section and the next are devoted, is found by solving (10) with neglect of the off-diagonal elements $C_{m,n}$. This approximation holds for atomic states which are not easily polarized or quenched, i.e., for transitions between two isolated atomic levels. The correction then comes merely from the fact that the atomic transition takes place in the field of the electron. In view of this, it is not surprising that the results we obtain involve quantities also encountered in the theory of elastic scattering of electrons.

For the two-state atom, then, with neglect of quenching and polarization,

$$i\hbar\dot{d}_\lambda=\sum_r\sum_\mu{}^0 b_{r\mu}J_r^*\exp(i\omega t+i\Omega_{\lambda\mu}t)\int\phi_\lambda{}^{1*}\phi_\mu{}^0 d\mathbf{R}, \quad (11)$$

$$i\hbar\dot{b}_{r\mu}=\sum_{\lambda'}{}^1 d_{\lambda'}J_r\exp(-i\omega t-i\Omega_{\lambda'\mu}t)\int\phi_\mu{}^{0*}\phi_{\lambda'}{}^1 d\mathbf{R}, \quad (12)$$

with $\hbar\Omega_{\lambda'\mu}=\epsilon_{\lambda'}{}^1-\epsilon_\mu{}^0$; the index r refers again to a photon of frequency $\omega_r=\omega_{\text{atom}}-\omega$, and the amplitudes d and b designate "dark" and "bright" states.

We now establish the fact that the exponential decay approximation for the initially excited atom is valid and independent of the electron state. From (12) we find an integral for $b_{r\mu}(t)$ which we substitute into Eq. (11):

$$i\hbar\dot{d}_\lambda=\sum_r\sum_\mu{}^0\frac{|J_r|^2}{i\hbar}\sum_{\lambda'}{}^1\int_0^t d_{\lambda'}$$

$$\times\exp[-i\omega(\tau-t)-i\Omega_{\lambda'\mu}(\tau-t)]d\tau\exp(i\Omega_{\lambda\lambda'}t)$$

$$\times\left[\int\phi_\mu{}^{0*}\phi_{\lambda'}{}^1 d\mathbf{R}\cdot\int\phi_\lambda{}^{1*}\phi_\mu{}^0 d\mathbf{R}\right]. \quad (13)$$

The assertion is: $d_{\lambda'}=a_{\lambda'}{}^1e^{-\gamma t}$, where $a_{\lambda'}{}^1$ is a constant. The proof is to note that it reduces the differential Eq. (13) for d_λ to an identity. The sum over oscillator states in (13) can be replaced by an integral.

$$\sum_r\rightarrow\int_0^\infty\rho_r d\omega_r\cong\int_{-\infty}^{+\infty}\rho_r d\omega_r,$$

where the symbols for the integration over angles of emission and the sum over polarization are compressed into ρ_r. Since the integral, $\int\phi_\mu{}^{0*}\phi_{\lambda'}{}^1 d\mathbf{R}$, is small for large $\Omega_{\lambda'\mu}$, we can take $\rho_r|J_r|^2$ out of the integral over ω_r, understanding that it remains constant over the spectral line, and replace it by its value $\rho|J|^2$ at $\omega=0$. Integrating over $x\equiv-(\omega+\Omega_{\lambda'\mu})$ and then over τ, we

obtain

$$\int_0^t d_{\lambda'}(\tau)\int_{-\infty}^{+\infty} e^{ix(\tau-t)}dx d\tau = \int_0^t e^{-\gamma(\tau-t)}2\pi\delta(\tau-t)d\tau a_{\lambda'}{}^1 e^{-\gamma t}=\pi a_{\lambda'}{}^1 e^{-\gamma t}.$$

Returning to (13), we find

$$\dot{d}_\lambda = -\pi\rho|J/\hbar|^2\sum_{\lambda'}{}^1 d_{\lambda'}\exp(i\Omega_{\lambda\lambda'}t) \times\sum_\mu{}^0\int\phi_\mu{}^{0*}\phi_{\lambda'}{}^1 d\mathbf{R}'\int\phi_\lambda{}^{1*}\phi_\mu{}^0 d\mathbf{R}$$

$$=-\gamma\sum_{\lambda'}{}^1 d_{\lambda'}\exp(i\Omega_{\lambda\lambda'}t)=-\gamma d_\lambda, \tag{14}$$

with $\gamma\equiv\pi\rho|J/\hbar|^2$. In establishing Eq. (14) the closure relation of the functions $\phi_\mu{}^0$ is used. This proves the exponential dependence of d_λ on time.

Equation (12) can be written

$$i\hbar\dot{b}_r\equiv\sum_\mu i\hbar\dot{b}_{r\mu}=J_r e^{(-\gamma-i\omega)t}\int\phi_0{}^*\phi_1 d\mathbf{R}, \tag{15}$$

provided we define

$$\phi_1\equiv\sum_\lambda{}^1 a_\lambda{}^1\phi_\lambda{}^1\exp(-i\epsilon_\lambda{}^1 t/\hbar),$$
$$\phi_0\equiv\sum_\mu{}^0\phi_\mu{}^0\exp(-i\epsilon_\mu{}^0 t/\hbar). \tag{16}$$

Both ϕ_0 and ϕ_1 represent electron wave packets moving in the Hartree field of the corresponding atomic state; in view of Eq. (9) they satisfy

$$(H_e+C_{n,n})\phi_n=i\hbar(\partial/\partial t)\phi_n. \tag{17}$$

The function ϕ_0 is, of course, not normalized.

Suppose the electron is initially in a single stationary state (λ), so that $a_{\lambda'}{}^1=\delta_{\lambda'\lambda}$. Then we find from (12) for the line intensity

$$I_U(\omega)=\rho_r\sum_\mu|b_{r\mu}(\infty)|^2 =\rho_r|J_r/\hbar|^2\sum_\mu\left|\int\phi_\mu{}^{0*}\phi_\lambda{}^1 d\mathbf{R}\right|^2\Big/\gamma^2+(\omega+\Omega_{\lambda\mu})^2. \tag{18}$$

A theory based on a formula of this type was developed by Rudkjobing[10]; his paper contains applications to the H and K doublets of Ca^+ and to the $4S$–$4P$ doublet of Si^{+++} at 4100 A.

We note that the normalization of $I_U(\omega)$ is maintained $(\int I_U d\omega=\int I d\omega)$, since

$$\int_{-\infty}^{+\infty}\frac{d\omega}{\gamma^2+(\omega+\Omega_{\lambda\mu})^2}\cong\int_{-\infty}^{+\infty}\frac{d\omega}{\gamma^2+\omega^2}$$

is independent of μ and, by the closure property,

$$\sum_\mu\left|\int\phi_\mu{}^{0*}\phi_\lambda{}^1 d\mathbf{R}\right|^2=1.$$

To evaluate this broadening quantitatively, we expand ϕ in eigenfunctions (u_λ) that are independent of the atomic state. In place of Eq. (16), we write

(using a slightly modified definition of d and b)

$$\phi_1=\sum_\lambda d_\lambda u_\lambda\exp(-i\epsilon_\lambda t/\hbar)$$

and

$$\phi_0=\sum_\lambda b_\lambda u_\lambda\exp(-i\epsilon_\lambda t/\hbar), \tag{19}$$

where $H_e u_\lambda=\epsilon_\lambda u_\lambda$. Equation (17) then leads to

$$i\hbar\dot{d}_\lambda=\sum_\mu(C_{1,1})_{\lambda,\mu}\exp[i(\epsilon_\lambda-\epsilon_\mu)t/\hbar]d_\mu\equiv\sum_\mu i\hbar p_{\lambda\mu}d_\mu, \tag{20}$$

and

$$i\hbar\dot{b}_\lambda=\sum_\mu(C_{0,0})_{\lambda,\mu}\exp[i(\epsilon_\lambda-\epsilon_\mu)t/\hbar]b_\mu\equiv\sum_\mu i\hbar q_{\lambda\mu}b_\mu; \tag{21}$$

with

$$(C)_{\lambda,\mu}=\int u_\lambda{}^*(C)u_\mu d\mathbf{R};$$

the p and q matrices are defined by (20) and (21).

Then

$$\int\phi_0{}^*\phi_1 d\mathbf{R}=\sum_\mu b_\mu{}^* d_\mu\equiv b(t)^\dagger d(t), \tag{22}$$

where b and d are column vectors. We now introduce time-development matrices T and U, defined by $d(t)=T(t)d(0)$, $b(t)=U(t)b(0)$; $T(0)=U(0)=E$. By (20) and (21) they satisfy $\dot{T}=pT$ and $\dot{U}=qU$. Since p and q are anti-Hermitean, T and U are unitary. Equation (22) then becomes

$$\int\phi_0{}^*\phi_1 d\mathbf{R}=b^\dagger d=b(0)^\dagger U^\dagger T d(0).$$

With exponential decay of the initially excited state Eq. (15) takes the form

$$i\hbar\dot{b}_r=J_r b(0)^\dagger U^\dagger T d(0)e^{-(\gamma+i\omega)t}. \tag{23}$$

For the broadened line intensity we then have,

$$I_U=\rho_r\sum_\mu|b_{r\mu}(\infty)|^2=\rho_r|\sum_\mu b_{r\mu}(\infty)|^2=\rho_r|b_r(\infty)|^2 =\rho_r|J_r/\hbar|^2\,\mathrm{Tr}\left[\rho^0\int_0^\infty d\tau T^\dagger U e^{-(\gamma-i\omega)\tau}\times\int_0^\infty d\tau' U^\dagger T e^{-(\gamma+i\omega)\tau'}\right], \tag{24}$$

where

$$\rho^0=d(0)d(0)^\dagger\quad\text{and}\quad b(0)b(0)^\dagger=E,$$

a unit matrix. Several steps in the preceding development require comment. First, we prove the equality of the sum of squares and the squared sum of the $b_{r\mu}$. It holds if

$$\sum_{\substack{\mu,\nu\\ \mu\neq\nu}} b_{r\mu}{}^* b_{r\nu}=0.$$

This is not true in general; e.g., when there are correlations between the phases of the individual $b_{r\mu}$. But in the absence of correlations every term $b_{r\mu}{}^* b_{r\nu}$ will vanish when supplied with the phase factor $\exp i(\alpha_\mu-\alpha_\nu)$ and averaged over some ensemble of phases.

The relation

$$b(0)b(0)^\dagger=E$$

rests on a similar argument. From (16) and (19) at

[10] M. Rudkjobing, Ann. astrophys. **12**, 229 (1949).

$t=0$,

$$\phi_0=\sum_\mu \phi_\mu{}^0=\sum_\lambda b_\lambda(0)u_\lambda.$$

It follows that

$$b_\lambda(0)=\sum_\mu\int u_\lambda{}^*\phi_\mu{}^0 d\mathbf{R};$$

and

$$b_\nu{}^*(0)b_\lambda(0)=\left(\sum_{\mu'}\int u_\nu\phi_{\mu'}{}^{0*}d\mathbf{R}'\right)\left(\sum_\mu\int u_\lambda{}^*\phi_\mu{}^0 d\mathbf{R}\right)$$

$$=\sum_\mu\int u_\nu\phi_{\mu'}{}^{0*}d\mathbf{R}'\int u_\lambda{}^*\phi_\mu{}^0 d\mathbf{R}=\delta_{\nu\lambda};$$

if cross-terms containing $\phi_{\mu'}{}^{0*}\phi_\mu{}^0$ with factors

$$\exp[i(\alpha_\mu-\alpha_{\mu'})]$$

are assumed to vanish on being averaged over phases.

We now transform

$$d_\mu\rightarrow d_\mu\exp[-i(C_{1,1})_{\mu\mu}t/\hbar]$$

and

$$b_\mu\rightarrow b_\mu\exp[-i(C_{0,0})_{\mu\mu}t/\hbar]$$

and thereby remove the diagonal elements from p and q.

For plane waves these diagonal matrix elements are independent of the electron state, since

$$|u_\lambda|^2\propto|\exp(i\mathbf{k}_\lambda\cdot\mathbf{R})|^2=1.$$

Designating such elements by

$$(C_{n,n})_{\lambda\lambda}\equiv C_{n,n},$$

we have

$$\int\phi_0{}^*\phi_1 d\mathbf{R}=\sum_\lambda b_\lambda{}^*d_\lambda\exp\{i[(C_{0,0})_{\lambda\lambda}-(C_{1,1})_{\lambda\lambda}]t/\hbar\}$$

$$=\exp\{i[C_{0,0}-C_{1,1}]t/\hbar\}\sum_\lambda b_\lambda{}^*d_\lambda. \quad (25)$$

In computing I_U from (23) and (24), we shall absorb the exponential factor occurring in (25) into $e^{-i\omega t}$ as a line shift. Its magnitude is the change in electrostatic interaction energy of the atomic states with a uniform charge distribution normalized to the free electron density.

Some interesting implications of Eq. (24) are easily established. First, if the perturbation elements $(C_{1,1})_{\lambda,\mu}$ and $(C_{0,0})_{\lambda,\mu}$ are equal, then no broadening occurs. This is reasonable, for in this case electron scattering by the atom is the same for the two atomic states; no change of the electron energy occurs with the radiative transition and no broadening is expected.

To demonstrate this result set $p(t)=q(t)$; then $T=U$; $T^\dagger U=E$; and

$$I_U=\rho_r|J_r/\hbar|^2\,\mathrm{Tr}\left\{\rho^0\int_0^\infty dte^{-(\gamma-i\omega)t}\int_0^\infty dte^{-(\gamma+i\omega)t}\right\}$$

$$=\frac{\rho_r|J_r/\hbar|^2}{\gamma^2+\omega^2},$$

since $\mathrm{Tr}(\rho^0)=1$.

Second, the normalization is maintained; i.e.

$$\int I_U d\omega_r=1.$$

This follows from (24) and a little algebra.

To prepare for the detailed work of the following section we now consider a specific expansion of the matrices T and U. As we shall see, this limits the range of validity of our results and restricts us to broadening effects comparable with the natural line width. We solve for T by iteration, writing

$$T(t)=E+T^{(1)}(t)+T^{(2)}(t)+\cdots,$$

where

$$T^{(n)}(t)=\int_0^t\int_0^{t_n}\cdots\int_0^{t_2}p(t_n)\cdots p(t_1)dt_1\cdots dt_{n-1}dt_n. \quad (26)$$

A similar expansion holds for U. This expansion is more strongly convergent for small C; i.e., little scattering. The first term, $T=U=E$, corresponds to no change in the perturber state and leaves the spectral line unchanged. Higher order terms are small if the interaction induces little change in the perturber during the lifetime of the excited atom. These higher terms give a correction to the spectral line which we evaluate to the second order in the elements of C.

On expanding,

$$I_U=\rho_r|J_r/\hbar|^2\,\mathrm{Tr}\Big\{\rho^0\Big[\int_0^\infty dte^{-a^*t}[E+(U^{(1)}+T^{\dagger(1)})$$

$$+(U^{(2)}+T^{\dagger(2)}+T^{\dagger(1)}U^{(1)})+\cdots]$$

$$\times\int_0^\infty dte^{-at}[E+(U^{\dagger(1)}+T^{(1)})$$

$$+(U^{\dagger(2)}+T^{(2)}+U^{\dagger(1)}T^{(1)})+\cdots]\Big]\Big\},$$

with $a\equiv\gamma+i\omega$.

Neglecting terms of order C^3 and higher, we find

$$I_U=\rho_r|J_r/\hbar|^2\,\mathrm{Tr}\Big\{\rho^0\Big[\frac{E}{|a|^2}+\frac{1}{a^*}\int_0^\infty dte^{-at}(U^{\dagger(1)}+T^{(1)})$$

$$+\frac{1}{a}\int_0^\infty dte^{-a^*t}(U^{(1)}+T^{\dagger(1)})$$

$$+\int_0^\infty dte^{-a^*t}(U^{(1)}+T^{\dagger(1)})\int_0^\infty dte^{-at}(U^{\dagger(1)}+T^{(1)})$$

$$+\frac{1}{a^*}\int_0^\infty dte^{-at}(U^{\dagger(2)}+T^{(2)}+U^{\dagger(1)}T^{(1)})$$

$$+\frac{1}{a}\int_0^\infty dte^{-a^*t}(U^{(2)}+T^{\dagger(2)}+T^{\dagger(1)}U^{(1)})\Big]\Big\}.$$

In practice, the statistical matrix ρ^0 is diagonal because the random phases of the d_μ, result in the cancellation of nondiagonal $d_\mu d_\nu^*$. Random phases imply a lack of interference patterns or other periodicities in the perturber current. Thus, on expanding the trace,

$$I_U=\rho_r|J_r/\hbar|^2\Big\{\frac{1}{|a|^2}+\frac{1}{a^*}\sum_\lambda \rho_{\lambda\lambda}{}^0\Big[\int_0^\infty dt e^{-at}(U_{\lambda\lambda}{}^{(1)*}+T_{\lambda\lambda}{}^{(1)})+\frac{1}{a}\int_0^\infty dt e^{-a^*t}(U_{\lambda\lambda}{}^{(1)}+T_{\lambda\lambda}{}^{(1)*})$$
$$+\sum_\mu\int_0^\infty dt e^{-a^*t}(U_{\lambda\mu}{}^{(1)}+T_{\mu\lambda}{}^{(1)*})\int_0^\infty dt e^{-at}(U_{\lambda\mu}{}^{(1)*}+T_{\mu\lambda}{}^{(1)})+\frac{1}{a}\int_0^\infty dt e^{-a^*t}(U_{\lambda\lambda}{}^{(2)}+T_{\lambda\lambda}{}^{(2)*}+\sum_\mu T_{\mu\lambda}{}^{(1)*}U_{\mu\lambda}{}^{(1)})$$
$$+\frac{1}{a^*}\int_0^\infty dt e^{-at}(U_{\lambda\lambda}{}^{(2)*}+T_{\lambda\lambda}{}^{(2)}+\sum_\mu U_{\mu\lambda}{}^{(1)*}T_{\mu\lambda}{}^{(1)})\Big]\Big\}.$$

From their defining Eq. (26), we find for the elements in the expansion of T:

$$T_{\lambda\lambda}{}^{(1)}=0,\quad T_{\lambda\mu}{}^{(1)}=p_{\lambda\mu}{}^0\frac{\exp(i\Omega_{\lambda\mu}t)-1}{i\Omega_{\lambda\mu}},\quad T_{\lambda\mu}{}^{(2)}=\sum_\nu\frac{p_{\lambda\nu}{}^0p_{\nu\mu}{}^0}{i\Omega_{\nu\mu}}\Big[\frac{\exp(i\Omega_{\lambda\mu}t)-1}{i\Omega_{\lambda\mu}}-\frac{\exp(i\Omega_{\lambda\nu}t)-1}{i\Omega_{\lambda\nu}}\Big];$$

where $p_{\mu\lambda}{}^0\equiv(C_{1,1})_{\mu\lambda}/i\hbar$, $q_{\mu\lambda}{}^0\equiv(C_{0,0})_{\mu\lambda}/i\hbar$. Because of these relations,

$$I_U(\omega)=\frac{\rho_r|J_r/\hbar|^2}{\gamma^2+\omega^2}\Big\{1+\sum_\lambda\rho_{\lambda\lambda}{}^0\sum_\mu\Big[\frac{|p_{\lambda\mu}{}^0-q_{\lambda\mu}{}^0|^2}{\gamma^2+(\omega+\Omega_{\lambda\mu})^2}+\frac{2}{\Omega_{\lambda\mu}{}^2}\Big\{(|q_{\lambda\mu}{}^0|^2-|p_{\lambda\mu}{}^0|^2)\frac{\omega\Omega_{\lambda\mu}}{\gamma^2+\omega^2}-|q_{\lambda\mu}{}^0-p_{\lambda\mu}{}^0|^2$$
$$+|q_{\lambda\mu}{}^0|^2\Big(\frac{\gamma^2+\omega^2+\omega\Omega_{\lambda\mu}}{\gamma^2+(\omega+\Omega_{\lambda\mu})^2}\Big)+|p_{\lambda\mu}{}^0|^2\Big(\frac{\gamma^2+\omega^2-\omega\Omega_{\lambda\mu}}{\gamma^2+(\omega-\Omega_{\lambda\mu})^2}\Big)$$
$$-\mathrm{Real}\Big(p_{\lambda\mu}{}^0q_{\lambda\mu}{}^{0*}\Big[\frac{\gamma^2+\omega^2+i\gamma\Omega_{\lambda\mu}+\omega\Omega_{\lambda\mu}}{\gamma^2+(\omega+\Omega_{\lambda\mu})^2}+\frac{\gamma^2+\omega^2-i\gamma\Omega_{\lambda\mu}-\omega\Omega_{\lambda\mu}}{\gamma^2+(\omega-\Omega_{\lambda\mu})^2}\Big]\Big)\Big\}\Big]\Big\}. \quad (27)$$

The frequency integral of the correction to the natural line $I_U-\rho_r|J_r/\hbar|^2/(\gamma^2+\omega^2)$ is zero. Thus the line intensity is still normalized; there is no quenching of the radiation.

When the matrix elements are real, I_U has a simpler form. Since the matrix elements considered in the next section are all real we avail ourselves of this advantage. On putting

$$p_{\lambda\mu}{}^0=-iP_{\lambda\mu}=-iP_{\mu\lambda},\quad q_{\lambda\mu}{}^0=-iQ_{\lambda\mu}=-iQ_{\mu\lambda},$$

Eq. (27) reduces to[11]

$$I_U(\omega)=\frac{\rho_r|J_r/\hbar|^2}{\gamma^2+\omega^2}\Big\{1+\frac{1}{\gamma^2+\omega^2}\sum_\lambda\rho_{\lambda\lambda}{}^0\sum_\mu\Big(\frac{P_{\lambda\mu}-Q_{\lambda\mu}}{\gamma^2+(\omega-\Omega_{\lambda\mu})^2}[2P_{\lambda\mu}(\omega^2-\gamma^2-\omega\Omega_{\lambda\mu})]$$
$$+\frac{P_{\lambda\mu}-Q_{\lambda\mu}}{\gamma^2+(\omega+\Omega_{\lambda\mu})^2}[P_{\lambda\mu}(\gamma^2+\omega^2)-Q_{\lambda\mu}(3\omega^2-\gamma^2+2\omega\Omega_{\lambda\mu})]\Big)\Big\}. \quad (28)$$

III. EVALUATION OF ELECTRON BROADENING FOR A TWO-STATE ATOM

Redistribution Factor

The present section, being restricted to effects that arise when the atom has but one upper state, is less interesting from the point of view of practical application than its sequels. For the restriction excludes the possibility of degeneracy, and therefore the occurrence of "linear Stark effects." These are numerically larger than the results computed here. The latter, however, are universal, independent of the accident of degeneracy and will therefore be considered first.

[11] S. Bloom, dissertation, Yale University, 1952 (unpublished).

The broadened intensity distribution given by II-(28) has the form

$$I_U=\frac{\rho|J/\hbar|^2}{\gamma^2+\omega^2}\Big\{1+\frac{R}{\gamma^2+\omega^2}\Big\}. \quad (1)$$

We shall refer to R as the *redistribution factor*. It is

$$R=\sum_\lambda\rho_{\lambda\lambda}{}^0\sum_\mu R_{\lambda\mu}, \quad (2)$$

and depends on the collision matrix elements between the initial (λ) and final (μ) unperturbed electron states. $R_{\lambda\mu}$ contains two resonance terms:

$$R_{\lambda\mu}=\frac{A_{\lambda\mu}}{\gamma^2+(\omega-\Omega)^2}+\frac{B_{\lambda\mu}}{\gamma^2+(\omega+\Omega)^2}, \quad (3)$$

where

$$\hbar\Omega=\epsilon_\lambda-\epsilon_\mu$$

represents the perturber energy loss and $\hbar\omega=E_1-E_0-\hbar\omega_r$. The maximum of the first term occurs when $\omega_r-\omega_a=-\Omega$; i.e., when the excess of radiation energy over the atomic energy has a sign opposite to that of the energy lost by the perturber; this combination does not conserve energy. The second, however, corresponds to conservation, since the perturber gain equals the radiation loss. A and B depend on the elements of C;

$$A_{\lambda\mu}=(P-Q)2P(\omega^2-\gamma^2-\omega\Omega)$$

and

$$B_{\lambda\mu}=(P-Q)[P(\gamma^2+\omega^2)-Q(3\omega^2-\gamma^2+2\omega\Omega)],$$

where

$$P=(C_{1,1})_{\lambda\mu}/\hbar \quad \text{and} \quad Q=(C_{0,0})_{\lambda\mu}/\hbar.$$

In the summation over λ, the first term in (3) has a resonance at

$$\Omega_{\lambda\mu}=\omega,$$

which means

$$\epsilon_\lambda-\epsilon_\mu=\hbar\omega_a-\hbar\omega_r. \tag{4}$$

If $\omega_r>\omega_a$, the radiation has received more energy than the atom possessed (in its excited state, which was assumed to be realized at $t=0$). Hence the radiation should have drawn energy from the electron. According to (4), however, ϵ_μ must be $>\epsilon_\lambda$, which means the electron, too, has gained energy.

Conversely, if $\omega_r<\omega_a$, the radiation field did not receive all the energy the atom had. But the electron did not get the excess either, because now $\epsilon_\lambda>\epsilon_\mu$ according to (4), which means the electron lost energy. Therefore this first term of (3) represents contributions to the line width which violate conservation of energy.

Characteristically this term is small. At resonance ($\Omega_{\lambda\mu}=\omega$) with respect to μ, its coefficient, $\omega^2-\gamma^2-\omega\Omega_{\lambda\mu}$ is $-\gamma^2$; it would therefore vanish if we did not include the energy uncertainty resulting from the natural line width.

The second term has a "resonance" (with respect to the variable λ) at

$$\Omega_{\mu\lambda}=\omega. \tag{5}$$

Hence,

$$\epsilon_\mu-\epsilon_\lambda=\hbar\omega_a-\hbar\omega_r.$$

If $\omega_r>\omega_a$, i.e., if the radiation field receives more energy than the atom can supply, then $\epsilon_\lambda>\epsilon_\mu$; that is, the electron looses energy. The second term therefore represents the energetically proper transfer of energy from the electron to the radiation field.

The greatest term in the summation over μ, defined by (5), has the coefficient

$$P_{\lambda\mu}(\gamma^2+\omega^2)-Q_{\lambda\mu}(3\omega^2-\gamma^2+2\omega\Omega_{\lambda\mu})$$
$$=P_{\lambda\mu}(\omega^2+\gamma^2)-Q_{\lambda\mu}(\omega^2-\gamma^2).$$

Interestingly, it survives even if γ is neglected.

To calculate I_U for a specific and simple case, we consider the $2p$ and $1s$ states of hydrogen. The perturber density (n) defines the volume per electron $V\equiv n^{-1}$. Using $g=V/(2\pi)^3$, the number of states per unit volume in wave number space, we replace the sum over final states by an integral:

$$\sum_\mu \rightarrow \int g d\mathbf{k}_\mu.$$

The integrations over direction and magnitude of $\mathbf{k}_\mu$ can be separated, since A and B change slowly with $|\mathbf{k}_\mu|$.

We note that P and Q are functions of $K^2=k_\lambda^2+k_\mu^2-2k_\lambda k_\mu\cos\Theta$. Hence we write $A=A(k_\mu,\cos\Theta)$ and $B=B(k_\mu,\cos\Theta)$. Consequently,

$$\sum_\mu R_{\lambda\mu}\rightarrow 2\pi g\int_{-1}^{+1} d\cos\Theta\int_0^\infty k_\mu^2 dk_\mu\left\{\frac{A(k_\mu,\cos\Theta)}{\gamma^2+(\omega-\Omega)^2}+\frac{B(k_\mu,\cos\Theta)}{\gamma^2+(\omega+\Omega)^2}\right\}.$$

The integrand is large at $\omega=\pm\Omega$. The values of A and B at these peaks, i.e., where their respective denominators reach a minimum, will be called the "resonance" values of k_μ. These occur at

$$\frac{\hbar^2}{2m}k_{\mu\mp}^2=\mp\left(\hbar\omega_a-\hbar\omega_r\mp\frac{\hbar^2}{2m}k_\lambda^2\right),$$

where the $-$ $(+)$ sign is for $A(B)$. We now take advantage of the slow variation of A and B with k_μ and obtain

$$\int R_{\lambda\mu}g d\mathbf{k}_\mu=2\pi g\left\{-2\gamma^2\int_0^\pi P_-(P_--Q_-)\sin\Theta d\Theta k_{\mu-}^2 G_-\right.$$
$$+\left[(\gamma^2+\omega^2)\int_0^\pi (P_+-Q_+)P_+\sin\Theta d\Theta\right.$$
$$\left.\left.-(\omega^2-\gamma^2)\int_0^\pi (P_+-Q_+)Q_+\sin\Theta d\Theta\right]k_{\mu+}^2 G_+\right, \tag{6}$$

where

$$G_\pm=\int_0^\infty \frac{dk_\mu}{\gamma^2+(\omega\pm\Omega)^2}, \tag{7}$$

and the arguments $+$ $(-)$ indicate evaluation at the resonance value $k_{\mu+}(k_{\mu-})$ which satisfies $\omega+\Omega=0$ $(\omega-\Omega=0)$. This approximation cannot be used if, for example, the radiation gains more energy than the perturber possessed initially. For this case the conservation resonance, $\omega+\Omega=0$, is not satisfied by a positive k_μ. Similarly, for the nonconservative resonance, the radiation must be restricted to lose less than the initial perturber energy. Thus the present method is not valid in the far wings of the line where $|\hbar\omega|>\epsilon_\lambda=(\hbar^2/2m)k_\lambda^2$.

Direct evaluation of the integral in (7) yields

$$G_{\pm}=\left(\frac{2m}{\hbar}\right)^{\frac{1}{2}}\frac{\pi i}{4\gamma}\left[\frac{-1}{(-i\gamma+b)^{\frac{1}{2}}}+\frac{1}{(i\gamma+b)^{\frac{1}{2}}}\right],$$

where

$$b=-(\hbar/2m)k_\lambda^2\pm\omega.$$

Ordinarily $\gamma\ll\epsilon_\lambda/\hbar$, and we find on expanding in powers of γ/b:

$$G_{\pm}=\left(\frac{2m}{\hbar}\right)^{\frac{1}{2}}\frac{\pi i}{4\gamma}\frac{1}{b^{\frac{1}{2}}}\left[\mp\left(1+\frac{i\gamma}{2b}+\cdots\right)\pm\left(1-\frac{i\gamma}{2b}\cdots\right)\right]$$

$$=\left(\frac{2m}{\hbar}\right)^{\frac{1}{2}}\frac{\pi i}{4\gamma}\frac{1}{b^{\frac{1}{2}}}\begin{cases}\pm 2\\ \pm i\gamma/b,\end{cases}\tag{8}$$

where the criterion of a real and positive G may be used to select the correct value. This will be done later [see Eq. (21)].

The electrostatic interaction is

$$C_{n,n}=-\frac{e^2}{|\mathbf{R}|}+e^2\int\frac{|\psi_n(\mathbf{r})|^2d\mathbf{r}}{|\mathbf{R}-\mathbf{r}|},$$

provided $\mathbf{r}$ and $\mathbf{R}$ represent the positions of atomic and perturbing electrons relative to the nucleus and $\psi_n(\mathbf{r})$ is the atomic wave function. The unperturbed electron functions are plane waves:

$$u_\lambda=(1/V^{\frac{1}{2}})\exp(i\mathbf{k}_\lambda\cdot\mathbf{R}).$$

Then the matrix elements between perturber states contain the factor

$$Z=\int\frac{e^{i\mathbf{K}\cdot\mathbf{R}}}{|\mathbf{R}-\mathbf{r}|}d\mathbf{R}=e^{i\mathbf{K}\cdot\mathbf{r}}\int\frac{\exp(i\mathbf{K}\cdot\mathbf{z})}{|\mathbf{z}|}d\mathbf{z}$$

$$=2\pi e^{i\mathbf{K}\cdot\mathbf{r}}\int_0^\infty z^2dz\int_0^\pi\frac{e^{iKz\cos\theta}}{z}\sin\theta d\theta$$

$$=\frac{4\pi}{K}e^{i\mathbf{K}\cdot\mathbf{r}}\int_0^\infty\sin Kzdz,\tag{9}$$

where $\mathbf{z}=\mathbf{R}-\mathbf{r}$ and $\mathbf{K}=\mathbf{k}_\mu-\mathbf{k}_\lambda$, the momentum transfer vector. The indeterminate τ is evaluated by using the artifice of a convergence factor $e^{-\alpha z}$, which leads to

$$\lim_{\alpha\to 0}\int_0^\infty e^{-\alpha z}\sin Kzdz=\lim_{\alpha\to 0}\frac{K}{\alpha^2+K^2}=\frac{1}{K}.\tag{10}$$

Hence,

$$Z=(4\pi/K^2)e^{i\mathbf{K}\cdot\mathbf{r}}.\tag{11}$$

Using this result and neglecting exchange, we find

$$\frac{1}{V}\int e^{i\mathbf{K}\cdot\mathbf{R}}C_{n,n}d\mathbf{R}\equiv(C_{n,n})_{\lambda\mu}=\frac{1}{V}\frac{4\pi e^2}{K^2}[F_{n,n}-1],$$

where

$$F_{n,n}\equiv\int e^{i\mathbf{K}\cdot\mathbf{r}}|\psi_n(\mathbf{r})|^2d\mathbf{r}\tag{12}$$

represents the usual form factor.

The atomic state and the momentum transfer vector, $\mathbf{K}$, determine the form factor (12). Its value is independent of the direction of $\mathbf{K}$ for a spherically symmetric charge distribution, in which case polar axis can be oriented along $\mathbf{K}$ to simplify the evaluation. The general situation will be discussed in detail in Sec. VI, where the shortcomings of the present simplified treatment are removed. The angular dependence of the $2p$ states will not trouble us, since in the absence of external fields the probabilities for the degenerate states are equal and we may take to a good approximation.

$$F_2\equiv\tfrac{1}{3}\textstyle\sum_m F_{2pm,2pm}.$$

Although this average can be evaluated with the use of the radial function alone, we include for completeness the individual F_{2pm}.

The atomic functions are

$$\psi_{1s,0}=\frac{1}{\pi^{\frac{1}{2}}}\frac{1}{a^{3/2}}e^{-r/a},$$

$$\psi_{2p,0}=\frac{1}{4(2\pi)^{\frac{1}{2}}}\frac{r}{a^{5/2}}e^{-r/2a}\cos\theta,$$

$$\psi_{2p,\pm1}=\frac{1}{8\pi^{\frac{1}{2}}}\frac{r}{a^{5/2}}e^{-r/2a}\sin\theta e^{\pm i\phi},$$

a being the first Bohr radius. To compute the F's we consider the integral

$$a_{100}\equiv\int e^{\alpha r}e^{i\mathbf{K}\cdot\mathbf{r}}d\mathbf{r}=2\pi\int_0^\infty r^2dre^{\alpha r}\int_{-1}^{+1}e^{iKr\mu}d\mu$$

$$=-\frac{8\pi\alpha}{(\alpha^2+K^2)^2};\quad(\alpha<0).\tag{13}$$

In calculating F_{2p0}, one encounters

$$a_{210}\equiv\int r^2e^{\alpha r}\mu^2e^{i\mathbf{K}\cdot\mathbf{r}}d\mathbf{r}=-\frac{d^2}{dK^2}a_{100}$$

$$=-32\pi\alpha\left[\frac{\alpha^2-5K^2}{(\alpha^2+K^2)^4}\right].\tag{14}$$

Similarly,

$$a_{21\pm1}\equiv\int r^2(1-\mu^2)e^{\alpha r}e^{i\mathbf{K}\cdot\mathbf{r}}d\mathbf{r}$$

$$=-\frac{2}{K}\frac{d}{dK}a_{100}=-\frac{64\pi\alpha}{(\alpha^2+K^2)^3}.\tag{15}$$

Making use of these relations, we find

$$F_1 \equiv F_{100,100} = 1/(1+\tfrac{1}{4}x)^2, \tag{16}$$

$$F_{210,210} = (1-5x)/(1+x)^4, \tag{17}$$

and

$$F_{21\pm1,21\pm1} = 1/(1+x)^3, \tag{18}$$

where

$$x = a^2K^2, \quad F_2 \equiv \langle F_{21m,21m}\rangle_m = (1-x)/(1+x)^4, \tag{19}$$

where $\langle\ \rangle_m$ means the average over m.

We are now ready to evaluate Eq. (6). There are integrals over the scattering angle Θ of the type

$$L \equiv \int_0^\pi (P-Q)^2 \sin\Theta d\Theta.$$

Transformation from Θ to the variable $x = a^2K^2$ proceeds by the relation

$$K^2 = k_\lambda^2 + k_\mu^2 - 2k_\lambda k_\mu \cos\Theta.$$

Thus $dx = 2a^2k_\lambda k_\mu \sin\Theta d\Theta$, and

$$L = \frac{1}{V^2}\frac{(4\pi)^2e^4}{\hbar^2}\frac{a^2}{2k_\lambda k_\mu}\int_{x_-}^{x_+}\left(\frac{F_1-F_2}{x}\right)^2 dx, \tag{20}$$

where

$$x_+ = a^2(k_\lambda + k_\mu)^2$$

and

$$x_- = a^2(k_\lambda - k_\mu)^2.$$

For hydrogen, $a = \hbar^2/me^2$ and

$$L = \frac{2}{V^2}(h/m)^2\frac{1}{k_\lambda k_\mu}\int_{x_-}^{x_+}\left(\frac{F_1-F_2}{x}\right)^2 dx.$$

Equation (6) takes a simple form when ω is small, i.e., when the perturber energy changes very little in comparison with its initial value. This is indeed a meaningful approximation, as will later be shown. Here

$$k_{\mu\pm}^2 = k_\lambda^2 \pm (2m/\hbar)\omega \cong k_\lambda$$

and, in view of Eq. (8),

$$G_\pm = (m/\hbar)(1/\gamma k_\lambda). \tag{21}$$

Furthermore, when a Taylor expansion is used,

$$\int_{x_-}^{x_+}\frac{(F_1-F_2)^2}{x^2}dx \cong \int_0^{4a^2k_\lambda^2}\frac{(F_1-F_2)^2}{x^2}dx \equiv S_U. \tag{22}$$

Consequently, near the line center all terms in Eq. (6) become independent of the resonance condition. Equation (6) then reads

$$\int R_{\lambda\mu}g d\mathbf{k}_\mu = 2\pi g(\omega^2-\gamma^2)k_\lambda^2 G\int_0^\pi (P-Q)^2\sin\Theta d\Theta$$

$$= 2\pi g(\omega^2-\gamma^2)k_\lambda^2\frac{\pi m}{\hbar\gamma k_\lambda}\frac{2}{V^2}\left(\frac{h}{m}\right)^2\frac{1}{(k_\lambda)^2}$$

$$\times\int_0^{4a^2k_\lambda^2}\left(\frac{F_1-F_2}{x}\right)^2 dx \equiv f(\omega^2-\gamma^2)S_U,$$

and

$$I_U = \frac{\rho_r|J_r/\hbar|^2}{\omega^2+\gamma^2}\left\{1+fS_U\frac{\omega^2-\gamma^2}{\omega^2+\gamma^2}\right\}. \tag{23}$$

We now define symbols and analyze the result.

$$f = 2nv_\lambda\sigma_\lambda/\gamma;$$

$n = 1/V$, the electron density; v_λ is the electron velocity in the state λ and σ_λ the cross section associated with its de Broglie wavelength,

$$\sigma_\lambda = \pi(\hbar/mv_\lambda)^2;$$

the function f is a measure of the relative number of atoms struck during the lifetime of the excited state.

The redistribution factor, R, [see Eq. (1)] has the property of decreasing the intensity within the half maximum and increasing it without. It is symmetric in ω. A more detailed investigation which allows ω to be large shows that far from the line center the redistribution factor goes to zero and a very slight shift to the red is introduced.

For large values of fS_U our result indicates that the intensity at the line center can become negative. In this case (indeed whenever $fS_U > 0.4$) terms of higher order in the collision matrix elements must be considered.

Evaluation of S_U

Since

$$(F_1-F_2)/x = (16u^3+17u^2+21u+18)/u^4(3+u)^2$$

[see (16) and (19)], where $u = 1+x$, the integral of $[(F_1-F_2)/x]^2$ can be transformed into a linear combination of integrals of the form $\int du/[(3+u)^4u^m]$. We expand this integrand:

$$\frac{1}{(3+u)^4u^m} = \frac{1}{3^{m+2}}\left[\sum_{t=0}^{m+2}(-1)^t\binom{m+2}{t}W^{t-m}\frac{1}{(3+u)^2}\right],$$

where

$$W = u/(3+u).$$

Integrating each term, using

$$\int\frac{W^{t-m}}{(3+u)^2}du = \frac{1}{3}\int W^{t-m}dW = \begin{cases}\dfrac{1}{3}\dfrac{W^{t-m+1}}{t-m+1} & t-m \neq -1\\ \tfrac{1}{3}\ln W & t-m=-1\end{cases},$$

we find for the indefinite integral:

$$\int \frac{du}{(3+u)^4u^m}=\frac{1}{3^{m+3}}\left[\sum_{t=0}^{m+2}(-1)^t\binom{m+2}{t}\left\{\frac{W^{t-m+1}}{t-m+1}\right\}\right],$$

with the understanding that the quantity in the curly brace is to be replaced by $\ln W$ when $t-m=-1$. The limits are $u=1$ and $u=1+x_+$.

When terms with like powers of W are collected, there results

$$\int\left(\frac{F_1-F_2}{x}\right)^2 dx=0.197W^3-1.04W^2+2.44W$$
$$-0.799\ln W+0.452W^{-1}-0.198W^{-2}+0.0434W^{-3}$$
$$-2.52\times10^{-3}W^{-4}-4.11\times10^{-3}W^{-5}+0.914$$
$$\times10^{-3}W^{-6}-0.261\times10^{-3}W^{-7}+C,$$

C being the constant of integration. This integral between the limits of 0 and x_+, i.e., between $W=\frac{1}{4}$ and $W=(1+x_+)/(4+x_+)$, namely S_U, is plotted as a function of the initial electron energy $\epsilon_\lambda/(e^2/2a)=\frac{1}{4}x_+=a^2k_\lambda^2$ (Fig. 1). The integrals in equation (6), when written in their exact form (20) are seen to depend on the final electron energy as well. This dependence enters as the difference between k_λ and k_μ in the limits of integration; and from Fig. 1, we see that they are quite insensitive to changes in ϵ_μ of the order of the line width $\hbar\gamma\sim10^{-6}$ ev, i.e., changes in the abscissa of this amount leave the ordinate unchanged. It is this circumstance which justifies both the resonance approximation and the use of S_U in place of an integral with exact limits in the region near the line center.

Numerical values to be used, typical of the positive column of an arc discharge, are: $n=10^{14}/\text{cm}^3$ and $\epsilon_\lambda=(1/25)(e^2/2a)\sim\frac{1}{2}$ ev ($T\sim5000°\text{K}$). We treat all the electrons as if they had this energy and take $\rho_{\nu\nu}{}^0=\delta_{\lambda\nu}$. As we have shown, the initial direction does

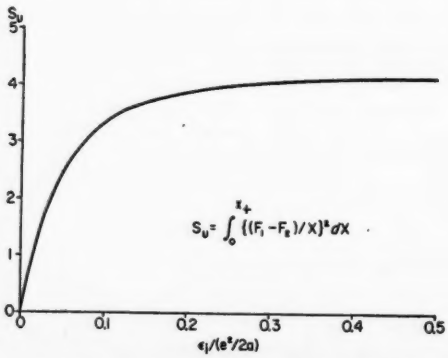


FIG. 1. S_U, the factor in universal broadening which depends on the atomic states, as a function of the initial electron energy, ϵ_λ.

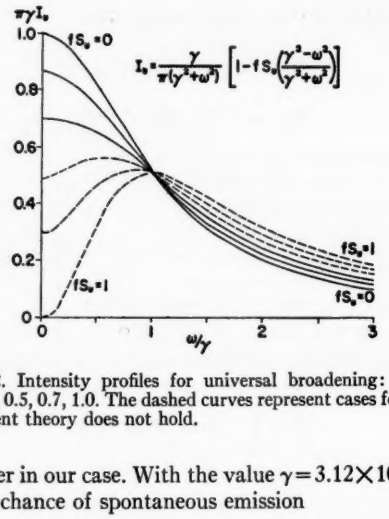


FIG. 2. Intensity profiles for universal broadening: $fS_U=0$, 0.13, 0.3, 0.5, 0.7, 1.0. The dashed curves represent cases for which the present theory does not hold.

not enter in our case. With the value $\gamma=3.12\times10^8/\text{sec}$[12] for the chance of spontaneous emission

$$I_U\cong\frac{\rho_r|J_r/\hbar|^2}{\gamma^2+\omega^2}\left\{1-0.13\left(\frac{\gamma^2-\omega^2}{\gamma^2+\omega^2}\right)\right\}. \tag{24}$$

This is plotted in Fig. 2. Curves for $fS_U=0.3$, 0.5, 0.7, and 1.0 corresponding to higher electron densities are also included in the graph.

The dotted curves, though given by our theory, are of course spurious because terms beyond the second order in the elements of C are important in these instances.

We note that the half-width of I_U [Eq. (23)] under the condition in which our approximation is valid (small fS_U) is given by

$$\gamma_{\frac{1}{2}}=\gamma(1+fS_U).$$

It is therefore suggested that we define the half-width for universal broadening as

$$\gamma_U\equiv fS_U\gamma.$$

Formula (23) can then be shown to be an approximation to first powers in γ_U/γ of

$$I_U=\frac{\rho_r|J_r/\hbar|^2(\gamma+\gamma_U)/\gamma}{\omega^2+(\gamma+\gamma_U)^2}, \tag{25}$$

which we expect to be valid even if γ_U/γ is large.

It may seem surprising that the natural line suffers almost no modification until fS_U reaches a value comparable to 1, at which point the natural line as such is destroyed rather suddenly. But this is merely another aspect of a point previously made: S_U is a number of order 1; $f\sim n\nu\sigma/\gamma$. Now t_e, the mean time between electron impacts with the atom, is $1/n\nu\sigma$, and t_a, the lifetime of the excited atomic state is of order $1/\gamma$. Hence $f\sim t_a/t_e$. If $f>1$, the electrons make many col-

[12] H. Bethe, *Handbuch der Physik* (Verlag Julius Springer, Berlin, 1933), second edition, Vol. 24, p. 444.

lisions during t_a, so that a static perturbation theory is applicable. The region of f under consideration seems to be characterized by the passage from a condition in which the time dependence matters to one in which a steady-state analysis is adequate.

The energy dependence of fS_U is easily obtained, because f is proportional to $(\epsilon_\lambda)^{-\frac{1}{2}}$ and $S_U(\epsilon_\lambda)$ has already been calculated. We find that fS_U is zero at both low and high energies because S_U and f vanish respectively at the extremes (Fig. 3).

In Fig. 4 we delimit the region of small electron broadening (of the universal kind here discussed) defined by $\gamma_U \leqslant \gamma/10$.

For negligible effects,

$$\frac{n}{10^{14}} < \frac{8.2}{S_U}\left(\frac{\epsilon_\lambda}{e^2/2a}\right)^{\frac{1}{2}}.$$

The theory developed here holds strictly in the region below the curve of Fig. 4. Immediately above it our theory is not adequate because of the neglect of higher powers of C. Far above this curve a statistical or time-independent type of treatment is appropriate.[13]

To remove the restriction to small n, distorted waves must be used to describe the electrons. This refinement is under consideration. The result is likely to be Eq. (25) with a γ_U not differing appreciably from that computed with plane waves.

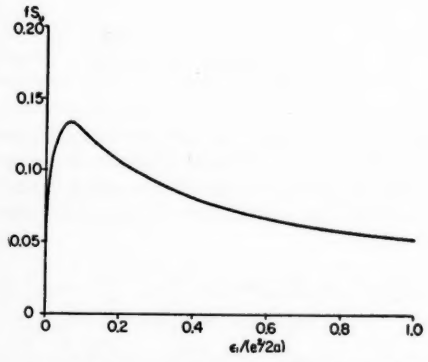


FIG. 3. fS_U, which is proportional to universal broadening, as a function of the initial electron energy, ϵ_i.

IV. INCLUSION OF POLARIZATION[14]

Polarization effects arise when the perturber is able to induce a dipole moment in the radiator or to orient an existing dipole moment. They depend, of course, on the speed and localizability of the perturbing ion or electron. Fixed ions give rise to the Stark shifted lines.

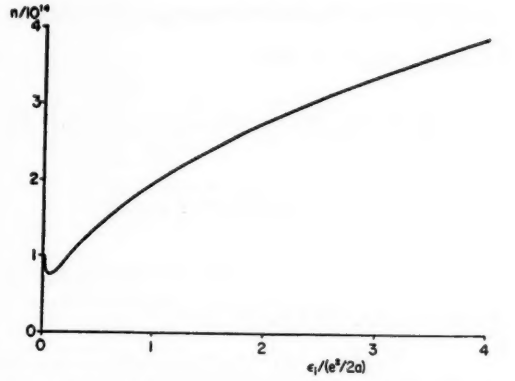


FIG. 4. Curve of density (n) *vs* electron energy (ϵ_i) delimiting the region of small n and large ϵ_i where universal broadening is negligible compared to the natural width (γ); the curve is the locus of points satisfying $fS_U = \gamma_U/\gamma = 1/10$.

As the perturber moves, the atom may or may not have time to be deformed or to re-orient itself. A measure of this failure to follow the changing field is given by the probability of nonadiabatic transitions. For very fast and frequent collisions, unpolarized, randomly oriented atoms are expected, and electron effects are negligible. We seek a criterion for the setting in of these polarization effects as the electron density increases and temperature decreases.

We study polarization effects as a correction to Holtsmark broadening and consider the Stark-shifted components of the first Lyman line from a hydrogen atom in the mean Holtsmark ion field, $F_{\text{ion}} = 3.26 e n_{\text{ion}}^{\frac{2}{3}}$. *Only three atomic states are considered:* $\psi_+ = (1/\sqrt{2}) \times (\psi_{2s,0} + \psi_{2p,0})$, $\psi_- = (1/\sqrt{2})(\psi_{2s,0} - \psi_{2p,0})$, and $\psi_{1s,0}$. The excited states correspond to dipoles ($3ea$) oriented along the z-axis of the atom which points at the ion. Their energies are displaced by $\pm 3eaF_{\text{ion}}$ from $E_2 \equiv -\frac{1}{4}(e^2/2a)$.

Deliberate exclusion of the states $\psi_{2p\pm1}$ is a very unrealistic procedure. Our reason is this. The two omitted states give rise to divergences requiring special treatment. Since we do not wish to encumber the basic analysis of the present section with mathematical difficulties, it seems indicated that the degenerate states, though present and far more effective in broadening the line, be first ignored.

Physically, a transition between ψ_+ and ψ_- corresponds to the re-orientation of a dipole, while a mixing-in of $\psi_{2p,+1}$ and $\psi_{2p,-1}$ represents induction of a dipole moment. In the present section, then, we make a somewhat artificial distinction between polarization by *re-orientation* and polarization by *induction*, and we treat here only the former effect.

The mathematical difficulties arising in connection with induction are removed in Sec. VII. This will prepare the way for the study of polarization by induction in Sec. VIII.

[13] H. Margenau and R. E. Meyerott (to be published).

[14] G. Wentzel, *Handbuch der Physik* (Verlag Julius Springer, Berlin, 1933), second edition, Vol. 24, p. 767. L. Spitzer, Phys. Rev. 55, 699 (1939).

The basic equation for the expansion coefficients is

$$i\hbar\dot{a}_{m\mu}\exp[-i(\epsilon_\mu{}^m+E_m)t/\hbar]$$
$$=\sum_n\sum_\lambda{}^n a_{n\lambda}J_{mn}\int\phi_\mu{}^{m*}\phi_\lambda{}^n d\mathbf{R}\exp[-i(\epsilon_\lambda{}^n+E_n)t/\hbar]$$
$$+\sum_{n\neq m}\sum_\lambda{}^n a_{n\lambda}\int\phi_\mu{}^{m*}C_{m,n}\phi_\lambda{}^n d\mathbf{R}$$
$$\times\exp[-i(\epsilon_\lambda{}^n+E_n)t/\hbar]. \quad (1)$$

The electron functions satisfy (II-(7))

$$[-(\hbar^2/2m)\nabla^2+C_{n,n}]\phi_\lambda{}^n=\epsilon_\lambda{}^n\phi_\lambda{}^n. \quad (2)$$

In order to neglect universal broadening we replace $C_{n,n}$ by a function which is independent of the atomic state; i.e.,

$$C_{n,n}\cong C_{m,m}\rightarrow C.$$

Below, where plane waves are used for the ϕ, we imply that $C\equiv 0$ in Eq. (2). As a consequence of this approximation,

$$\int\phi_\mu{}^{m*}\phi_\lambda{}^n d\mathbf{R}\cong\int\phi_\mu{}^*\phi_\lambda d\mathbf{R}=\delta_{\mu\lambda}, \quad (3)$$

with the equation

$$[-(\hbar^2/2m)\nabla^2+C]\phi_\mu=\epsilon_\mu\phi_\mu \quad (4)$$

determining the ϕ_μ. The part of the electron-atom interaction retained corresponds to the changes induced in the atom by the second sum in Eq. (1), and the effect of these nonradiative transitions will be called polarization broadening. To see how it enters, we start with the growth equations for the three-state atom. In fairly obvious notation,

$$i\hbar\dot{d}_{+\mu}\exp[-i(\epsilon_\mu+E_+)t/\hbar]$$
$$=\sum_r b_{r\mu}J_r{}^*\exp[-i(\epsilon_\mu+E_{1s}+h\nu_r)t/\hbar]$$
$$+\sum_\lambda d_{-\lambda}\int\phi_\mu{}^*C_{+,-}\phi_\lambda d\mathbf{R}$$
$$\times\exp[-i(\epsilon_\lambda+E_-)t/\hbar]; \quad (5)$$

$$i\hbar\dot{d}_{-\nu}\exp[-i(\epsilon_\nu+E_-)t/\hbar]$$
$$=\sum_r b_{r\nu}J_r{}^*\exp[-i(\epsilon_\nu+E_{1s}+h\nu_r)t/\hbar]$$
$$+\sum_\lambda d_{+\lambda}\int\phi_\nu{}^*C_{-,+}\phi_\lambda d\mathbf{R}$$
$$\times\exp[-i(\epsilon_\lambda+E_+)t/\hbar]; \quad (6)$$

$$i\hbar\dot{b}_{r\mu}=J_r\{d_{+\mu}\exp[i(E_{1s}+h\nu_r-E_+)t/\hbar]$$
$$+d_{-\mu}\exp[i(E_{1s}+h\nu_r-E_-)t/\hbar]\}; \quad (7)$$

and

$$C_{+,-}\equiv\int\psi_+{}^*C\psi_-d\tau.$$

We know that the effect of the terms with J in the equations for $\dot{d}$ can be approximated by an exponential decay; therefore Eqs. (5)–(7) become

$$i\hbar\dot{d}_{+\mu}=-i\hbar\gamma_+d_{+\mu}+\sum_\lambda d_{-\lambda}(C_{+,-})_{\mu\lambda}$$
$$\times\exp[i(\omega_{+,-}+\Omega_{\mu\lambda})t], \quad (8)$$

$$i\hbar\dot{d}_{-\nu}=-i\hbar\gamma_-d_{-\nu}+\sum_\lambda d_{+\lambda}(C_{-,+})_{\nu\lambda}$$
$$\times\exp[i(\omega_{-,+}+\Omega_{\nu\lambda})t], \quad (9)$$

$$i\hbar\dot{b}_{r\mu}=J_r[d_{+\mu}e^{-i\omega_+t}+d_{-\mu}e^{-i\omega_-t}]. \quad (10)$$

Here

$$\hbar\omega_{+,-}\equiv E_+-E_-,\quad \hbar\omega_+\equiv E_+-E_{1s}-h\nu_r,$$
$$\hbar\omega_-\equiv E_--E_{1s}-h\nu_r,$$

and

$$(C_{+,-})_{\mu\lambda}=\int\phi_\mu{}^*(C_{+,-})\phi_\lambda d\mathbf{R}.$$

To solve these equations we first put

$$d_{+\mu}\equiv D_{+\mu}e^{-\gamma_+t}\quad\text{and}\quad d_{-\mu}\equiv D_{-\mu}e^{-\gamma_-t}. \quad (11)$$

This reduces the growth equations to

$$i\hbar\dot{D}_{+\mu}=\sum_\lambda D_{-\lambda}(C_{+,-})_{\mu\lambda}$$
$$\times\exp[i(\omega_{+,-}+\Omega_{\mu\lambda})t-(\gamma_--\gamma_+)t], \quad (12)$$

$$i\hbar\dot{D}_{-\nu}=\sum_\lambda D_{+\lambda}(C_{-,+})_{\nu\lambda}$$
$$\times\exp[i(\omega_{-,+}+\Omega_{\nu\lambda})t+(\gamma_--\gamma_+)t], \quad (13)$$

$$i\hbar\dot{b}_{r\mu}=J_r\{D_{+\mu}\exp[-(\gamma_++i\omega_+)t]$$
$$+D_{-\mu}\exp[-(\gamma_-+i\omega_-)t]\}. \quad (14)$$

As the initial conditions we consider an excited atom in state ψ_+, no radiation present and the perturbing electron in the state

$$\phi_i{}^+\cong\frac{1}{\sqrt{V}}\exp(i\mathbf{k}_i\cdot\mathbf{R}-i\epsilon_it/\hbar). \quad (15)$$

Further, all amplitudes $d_{+\lambda}$, are assumed to remain small compared to d_{+i} during the lifetime of the excited state. That is, polarizing transitions which return the atom to the initial state before the atom radiates are unlikely. Then

$$i\hbar\dot{D}_{+i}=\sum_\mu(C_{+,-})_{i\mu}D_{-\mu}$$
$$\times\exp[i(\omega_{+,-}+\Omega_{i\mu})t-(\gamma_--\gamma_+)t]; \quad (16)$$

$$i\hbar\dot{D}_{-\mu}=\sum_\lambda(C_{-,+})_{\mu\lambda}D_{+\lambda}$$
$$\times\exp[i(\omega_{-,+}+\Omega_{\mu\lambda})t+(\gamma_--\gamma_+)t]$$
$$\cong(C_{-,+})_{\mu i}D_{+i}\exp[i(\omega_{-,+}+\Omega_{\mu i})t+(\gamma_--\gamma_+)t]. \quad (17)$$

From (14) and because cross-terms with random phase

factors cancel,

$$I(\omega)=\sum_\mu \rho_r |b_{r\mu}(t=\infty)|^2$$

$$=\rho_r |J_r/\hbar|^2\left\{\left|\int_0^\infty \exp[-(\gamma_+ + i\omega_+)t]D_{+i}dt\right|^2\right.$$

$$\left.+\sum_\mu\left|\int_0^\infty \exp[-(\gamma_- + i\omega_-)t]D_{-\mu}dt\right|^2\right\}. \quad (18)$$

Clearly, when all of the C matrix elements vanish, the D are constants: $D_{+i}=1$ and $D_{-\mu}=0$. Then $I(\omega)$, as expected, contains only the + component of the spectrum, and this line has the natural breadth.

It will now be shown that

$$D_{+i}\cong e^{-\gamma_P t}. \quad (19)$$

The spectral line is then broadened in accordance with the uncertainty principle, γ being replaced by $\gamma+\gamma_P$.

To find γ_P we substitute $D_{-\mu}$ found from (17) into (16):

$$\dot{D}_{+i}=-\gamma_P D_{+i}\cong -\frac{1}{\hbar^2}\sum_\mu |(C_{+,-})_{i\mu}|^2 \int_0^t D_{+i}$$

$$\times\exp\{[i(\omega_{-,+}+\Omega_{\mu i})+(\gamma_- -\gamma_+)](\tau-t)\}d\tau. \quad (20)$$

This result is general and free from the limitation to a single large d_{+i}. Had we retained all terms on the right of (17), then the place of (20) would be taken by

$$i\hbar\dot{D}_{+i}e^{i\alpha_i}=\sum_\mu\sum_\lambda (C_{+,-})_{i\mu}(C_{-,+})_{\mu\lambda}$$

$$\times\int_0^t D_{+\lambda}\exp(i\alpha_\lambda)\exp\{[i(\omega_{-,+}+\Omega_{\mu\lambda})$$

$$+(\gamma_- -\gamma_+)](\tau-t)\}d\tau\exp(i\Omega_{ii}t), \quad (21)$$

where the phase factors $\exp(i\alpha_\lambda)$ are written explicitly. When we average over random phases (21) reduces to (20), since the average of $\exp[i(\alpha_\lambda-\alpha_i)]$ is zero unless $\lambda=i$.

Before summing over μ we evaluate the matrix elements.

Preliminary account is taken of the random orientation of the direction of motion of the incident electron with respect to the atom by averaging over directions of the atomic z axis for a given $\mathbf{k}_i$. (See Sec. VI for an accurate treatment of the angular dependence of form factors.) Hence we replace $\psi_-{}^*\psi_+$ by the spherically symmetric

$$\langle\psi_-{}^*\psi_+\rangle_{Av}=\frac{1}{4\pi}\int_0^{2\pi}d\phi\int_0^\pi \sin\theta d\theta\psi_-{}^*\psi_+$$

$$=\tfrac{1}{4}\int_{-1}^{+1}[|\psi_{2s,0}|^2-|\psi_{2p,0}|^2]d\mu$$

$$=\tfrac{1}{2}[(R_{2s})^2-\tfrac{1}{3}(R_{2p})^2], \quad (22)$$

where

$$(R_{2p})^2\equiv\frac{r^2}{32\pi}\frac{e^{-r/a}}{a^5},\quad (R_{2s})^2\equiv\frac{[2-(r/a)]^2}{32\pi}\frac{e^{-r/a}}{a^3}.$$

Then, with the use of plane waves,

$$|(C_{+,-})_{\mu i}|^2$$

$$\cong\left|\int\int\frac{e^{i\mathbf{K}\cdot\mathbf{R}}}{V}\left(\frac{e^2}{|\mathbf{R}-\mathbf{r}|}-\frac{e^2}{|\mathbf{R}|}\right)\langle\psi_-{}^*\psi_+\rangle_{Av}d\mathbf{r}d\mathbf{R}\right|^2$$

$$=\left|\frac{4\pi e^2}{K^2}\cdot\tfrac{1}{2}\int\frac{e^{i\mathbf{K}\cdot\mathbf{r}}}{V}[(R_{2s})^2-\tfrac{1}{3}(R_{2p})^2]d\mathbf{r}\right|^2. \quad (23)$$

The interaction with the nucleus, e^2/R, contributes nothing to γ_P, since ψ_+ is orthogonal to ψ_-. The integral (23) over the atomic coordinates is independent of the direction of $\mathbf{K}$ which is along the polar axis;

$$\int e^{i\mathbf{K}\cdot\mathbf{r}}\{(R_{2s})^2-\tfrac{1}{3}(R_{2p})^2\}d\mathbf{r}$$

$$=2\pi\int_{-1}^{+1}d\mu\int_0^\infty dr e^{iKr\mu}\frac{[4-4(r/a)+\frac{2}{3}(r/a)^2]}{32\pi a^3}e^{-r/a}r^2$$

$$=\frac{1}{(1+x)^2}-\frac{(3-x)}{(1+x)^3}+\frac{2(1-x)}{(1+x)^4}=\frac{2x(x-1)}{(1+x)^4}. \quad (24)$$

Here we have made use of the following earlier results. The integral is considered as the sum of three integrals corresponding to the three terms in the square brace (24). The first is related to $F_{100,100}$ and is adapted from the integral a_{100} [see III-(13)]. The third is just $2F_2$ [see III-(19)]. The second was not previously encountered but is evaluated in a similar way and yields

$$-\frac{1}{4iKa}\left[\frac{2}{(-iKa-1)^3}-\frac{2}{(iKa-1)^3}\right]=-\frac{3-x}{(1+x)^3}. \quad (25)$$

Thus, from (23) and (24),

$$|(C_{+,-})_{i\mu}|^2=\left(\frac{4\pi e^2a^2}{V}\frac{(x-1)}{(x+1)^4}\right)^2, \quad (26)$$

which is a function of the momentum transfer through

$$x=a^2K^2=a^2(k_i{}^2+k_\mu{}^2-2k_ik_\mu\cos\Theta).$$

The evaluation of the sum over states μ in (20) can be simplified because the change in the matrix element, as $\epsilon_\mu=(\hbar^2/2m)k_\mu{}^2$ deviates from $\epsilon_i+\hbar\omega_{+,-}$ by $\hbar(\gamma_P+\gamma_+-\gamma_-)$, is very small. The time integral, however, peaks in this region of ϵ_μ. Hence we can factor the integrals over the magnitude and direction of $\mathbf{k}_\mu$; i.e., from (20)

$$\frac{dD_{+i}}{dt}=-\frac{1}{\hbar^2}\sum_\mu|(C_{+,-})_{i\mu}|^2\int_0^t\cdots d\tau$$

$$\cong\frac{-V}{(2\pi\hbar)^2}\int_0^\pi|(C_{+,-})_{i\mu_0}|^2k_{\mu 0}\sin\Theta d\Theta$$

$$\times\int_0^\infty\left[\int_0^t\cdots d\tau\right]k_\mu dk_\mu, \quad (27)$$

where (see Sec. III)

$$\Sigma_\mu \rightarrow \int \frac{V}{(2\pi)^3} d\mathbf{k}_\mu$$

and

$$(\hbar^2/2m)k_{\mu 0}^2 \equiv \hbar\omega_{+,-} + \epsilon_i.$$

By inverting the order of integration over τ and k_μ, replacing $k_\mu dk_\mu$ by $(m/\hbar)d(\omega_{-,+}+\Omega_{\mu i})$, and using

$$\int_{-\infty}^{+\infty} \exp[i(\omega_{-,+}+\Omega_{\mu i})(\tau-t)]d(\omega_{-,+}+\Omega_{\mu i}) = 2\pi\delta(\tau-t),$$

the last integral factor in (27) becomes

$$\int_0^t D_{+i}(\tau)\exp[+(\gamma_- - \gamma_+)(\tau-t)]\frac{2m\pi}{\hbar}\delta(\tau-t)d\tau = \frac{\pi m}{\hbar}D_{+i}(t). \quad (28)$$

Hence

$$\dot{D}_{+i} = -\gamma_P D_{+i} = -\frac{V}{(2\pi\hbar)^2}\int_0^\pi |(C_{+,-})_{i\mu 0}|^2 k_{\mu 0} \times \sin\Theta d\Theta \cdot \frac{\pi m}{\hbar} D_{+i}.$$

This leads to the result

$$\gamma_P = 2nv_i\sigma_i S_P, \quad (29)$$

with

$$S_P \equiv \int_0^{x_+}\left[\frac{(x-1)}{(1+x)^4}\right]^2 dx = \frac{11}{105} - \frac{1}{5(1+x_+)^5} + \frac{2}{3(1+x_+)^6} - \frac{4}{7(1+x_+)^7},$$

$$x_+ = a^2(k_i+k_\mu)^2 \cong 4(k_i a)^2 = 4\epsilon_i/(e^2/2a),$$

$$\sigma_i \equiv \pi(\hbar/mv_i)^2.$$

If we say, somewhat arbitrarily, that polarization by orientation can be neglected when

$$\gamma_P < \gamma/10, \quad (30)$$

there is defined region of low electron density and high energy delimited by

$$\frac{n}{10^{14}} = \left(\frac{\epsilon_i}{(e^2/2a)}\right)^{\frac{1}{2}}\frac{8.2}{S_P}. \quad (31)$$

But this is of theoretical interest only, for it will be seen that the orientation effect we have here considered contributes far less to the line width than the states thus far omitted. To return to our example: at $\epsilon_i \sim \frac{1}{2}$ ev, an electron density of the order $n \lesssim 10^{15}$ induces few reorientations of the (three-state) radiating atom and hence this polarization effect can be neglected. If $\epsilon_i = \frac{1}{2}$ ev and $n = 10^{14}$, then with $\gamma = 3.12\times10^8$, $S_P = 0.081$ and

$$\gamma_P/\gamma = fS_P \cong 0.0049. \quad (32)$$

This result is subject to modification because of the angular dependence of the form factors, which is discussed in Sec. VI. Other atomic states (especially $\psi_{2p,\pm1}$) are included later.

V. QUENCHING AND STARK BROADENING

A. Quenching

For quenching, the intensity distribution is the same as for polarization broadening:

$$I_Q(\omega) \propto [\omega^2 + (\gamma+\gamma_Q)^2]^{-1}, \quad (1)$$

provided γ_Q is equal to half the transition rate from the initial atomic level induced by electron impacts. That is,

$$\gamma_Q = \tfrac{1}{2}nv_iQ_i \quad (2)$$

for initial electron velocity v_i. Here Q_i is the quenching cross section for hydrogen,[15]

$$Q_{i\ (2p\rightarrow 1s)} = \int \frac{k_\lambda}{k_i}\left(\frac{m}{2\pi\hbar^2}\right)^2\left(\frac{4\pi e^2}{K^2}\right)^2 F_{1s,2p}{}^2 \sin\Theta d\Theta d\Phi. \quad (3)$$

Now k_λ is fixed by conservation of energy, and (see Appendix)

$$\gamma_Q = 2nv_i\sigma_i\int_{x_-}^{x_+}\frac{|F_{1s,2p}|^2}{x^2}dx = \gamma f_i S_Q. \quad (4)$$

The form factor is evaluated as before:

$$F_{100,210} = \frac{1}{4\sqrt{2}\pi a^4}\int e^{-3r/2a}\cos\theta r e^{i\mathbf{K}\cdot\mathbf{r}}d\mathbf{r} = 3iKa2^{15/2}/[9+4K^2a^2]^3. \quad (5)$$

Furthermore,

$$F_{100,21\pm1} = 0, \quad \text{since} \quad \int_0^{2\pi} e^{\pm i\phi}d\phi = 0.$$

This means that only $\frac{1}{3}$ of the atoms will be so oriented as to allow a quenching transition. Therefore one can use $F_{100,210}$ in (4) and multiply by the factor $\frac{1}{3}$.

The integral over x, $\int |F_{1s,2p}|^2/x^2 dx$, contains the factor

$$\int \frac{dx}{x(9+4x)^6} = -\frac{1}{9^6}\left[-\ln(S) + 5\times4S - \frac{10\times4^2}{2}S^2 + \frac{10\times4^3}{3}S^3 - \frac{5\times4^4}{4}S^4 + \frac{4}{5}S^5\right],$$

where

$$S = x/(9+4x). \quad (6)$$

[15] G. Wentzel, *Handbuch der Physik* (Verlag Julius Springer, Berlin, 1933), second edition, Vol. 24, p. 738.

For $\epsilon_i=(1/25)(e^2/2a)$ and $n=10^{14}$, $k_i=1/5a$, $S_Q=0.021$, and

$$\gamma_Q/\gamma=0.0013. \tag{7}$$

B. Stark Broadening by Ions[16]

Stark broadening by ions results from variations of the electric field intensity at the atom. The mean field for the ions given by Holtsmark is $F=3.26en_i^{\frac{2}{3}}$, where n_i is the ion density. This splits the degenerate hydrogen $n=2$ states into three components. The separation of the outer two, $6Fea$, is linear in F. Using this estimate the half-width of the broadened line, we find

$$\gamma_{\text{Stark}}=3.26en^{\frac{2}{3}}\times 3ea/\hbar=2.5\times 10^{10}/\text{sec} \tag{8}$$

for $n=10^{14}/\text{cm}^3$. Consequently, Stark broadening by ions is at least 100 times larger than the universal broadening.

A comment on the mass dependence of the broadening effect computed in this section is in order. The effect is expected to be small for heavy ions since it arises from elastic scattering collisions, and the ions are scattered less. This is indeed borne out by our formulas.

Broadening is proportional to $fS_U=2nv\sigma S_U/\gamma$. Let the ions have a reduced mass (with respect to the radiating atom) M. Now $v\propto M^{-\frac{1}{2}}$ for ions of the same energy, $\sigma=\pi\lambda^2\propto(\hbar/Mv)^2\propto M^{-1}$; hence $f\propto M^{-\frac{1}{2}}$. On the other hand, S_U depends on M through the upper limit of Eq. III-(22), and it increases with k_μ^2. At low energy its variation with k_μ^2 is approximately linear. However, $k_\mu^2\propto M$. In all, therefore, $fS_U\propto M^{-\frac{1}{2}}$.

Protons are thus less effective in broadening the line by a factor of at least 30.

VI. ANGULAR DEPENDENCES IN THE FORM FACTORS

In the preceding work, we have made much use of the form factor

$$F_{n,m}(\mathbf{K})=\int e^{i\mathbf{K}\cdot\mathbf{r}}\psi_n^*\psi_m d\mathbf{r}. \tag{1}$$

Throughout we simplified these integrals by taking the momentum transfer vector $\mathbf{K}$ to be directed along the z-axis of the atom. When the product of the atomic functions $\psi_n^*\psi_m$ is spherically symmetric, this is of no consequence. We now extend the calculation and demonstrate that the orientation of the atom with respect to the momentum transfer vector is generally unimportant for small Ka.

Let us evaluate

$$F_{2p0,\,2p0}=\int e^{i\mathbf{K}\cdot\mathbf{r}}|\psi_{2p,\,0}{}^{z'}|^2 d\mathbf{r}, \tag{2}$$

where the superscript z' specifies the direction of the polar axis for the wave function, $\psi_{2p,\,0}{}^{z'}$. The direction of $\mathbf{K}$ is the polar (z) axis of $\psi_{2pm}{}^{z}$. The angles Θ, Φ define the orientation of z' with respect to z. In terms of these angles, we expand

$$\begin{aligned}\psi_{2p,\,0}{}^{z'}&=R_{2p}\cos\Theta'\\&=R_{2p}[\cos\Theta\cos\theta+\sin\Theta\sin\theta\cos(\phi-\Phi)]\\&=\cos\Theta\psi_{2p,\,0}{}^{z}+(\sin\Theta/\sqrt{2})[e^{-i\Phi}\psi_{2p,+1}{}^{z}+e^{i\Phi}\psi_{2p,-1}{}^{z}],\end{aligned} \tag{3}$$

with

$$R_{2p}\equiv\frac{1}{4(2\pi)^{\frac{1}{2}}}\frac{r}{a^{5/2}}e^{-r/2a}.$$

Furthermore,[17]

$$e^{iKz}=\sum_L(2L+1)i^L P_L(\cos\theta)f_L(Kr), \tag{4}$$

where

$$f_L(r)=(\pi/2Kr)^{\frac{1}{2}}J_{L+1/2}(Kr).$$

Integration over Φ annuls all terms in $F_{2p0,\,2p0}$ which contain Φ-dependent factors. Hence the integral reduces to

$$\begin{aligned}F_{2p0,\,2p0}=2\pi\int_{-1}^{+1}d\mu\int_0^\infty r^2dre^{iKz}\{&[\cos^2\Theta|\psi_{2p,\,0}(z)|^2\\&+\tfrac{1}{2}\sin^2\Theta(|\psi_{2p,+1}(z)|^2+|\psi_{2p,-1}(z)|^2)]\}\\&+2\pi\int_{-1}^{+1}d\mu\int_0^\infty r^2dr(R_{2p})^2e^{iKz}\\&\times\{\cos^2\Theta\cos^2\theta+\tfrac{1}{2}\sin^2\Theta\sin^2\theta\}.\end{aligned} \tag{5}$$

When we replace e^{iKz} by its expansion, most of the terms contribute nothing because of the orthogonality of the Legendre polynomials. On using well-known relations[18] between these polynomials, we find

$$\begin{aligned}F_{2p,\,0;\,2p,\,0}&=2\pi\int_0^\infty r^2dr(R_{2p})^2\{\cos^2\Theta[\tfrac{2}{3}f_0-(4/3)f_2]\\&\qquad+\tfrac{1}{2}\sin^2\Theta[(4/3)f_0+(4/3)f_2]\}\\&=\frac{4\pi}{3}\int_0^\infty r^2dr(R_{2p})^2\\&\qquad\times\{f_0+f_2[\sin^2\Theta-2\cos^2\Theta]\}.\end{aligned} \tag{6}$$

The radial integral for f_0,

$$\frac{1}{a^5}\int_0^\infty\frac{\sin Kr}{Kr}r^4e^{-r/a}dr=\frac{24(1-x)}{(1+x)^4}, \tag{7}$$

where $x\equiv(Ka)^2$. Further,[19]

$$\frac{1}{a^5}\int_0^\infty r^4e^{-r/a}f_2(Kr)dr=\frac{48x}{(1+x)^4}. \tag{8}$$

[16] For a discussion of the quantum theory of broadening by ions see reference 13.

[17] N. F. Mott and H. S. W. Massey, *Atomic Collisions* (Clarendon Press, Oxford, 1949), second edition, p. 22.

[18] H. Margenau and G. M. Murphy, *Mathematics of Physics and Chemistry* (D. Van Nostrand Company, Inc., New York, 1943), Chap. 3, Secs. 4 and 5.

[19] See reference 18, Chap. 3, Sec. 8.

Combining these results, we find

$$F_{2p,0;2p,0}=\frac{(1-x)}{(1+x)^4}+[\sin^2\Theta-2\cos^2\Theta]\frac{2x}{(1+x)^4}.$$

Clearly, at $\Theta=0$ we have

$$F_{2p,0;2p,0}=(1-5x)/(1+x)^4, \tag{9}$$

which is the result found directly for $\mathbf{K}$ along the (z') polar axis of the atom [III-(17)]. Further, if we average over random angles of collisions (i.e., over Θ) the term in brackets vanishes, and the result is the quantity used in our earlier study [III-(19)]:

$$\langle F_{2p0,2p0}\rangle_\Theta=(1-x)/(1+x)^4=F_2. \tag{10}$$

To proceed more rigorously, we use the sum over initial states to average over orientations of collisions (i.e., $\sum_i \rho_{ii}^0\to\frac{1}{2}\int_0^\pi \sin\Theta d\Theta$). Let us put

$$\tfrac{1}{2}\int_0^\pi [F_1-F_{2p0,2p0}(\Theta)]^2\sin\Theta d\Theta=(F_1-F_2)^2+\Delta, \tag{11}$$

so that the deviation from our earlier result is

$$\Delta\equiv\left[\frac{2x}{(1+x)^4}\right]^2\tfrac{1}{2}\int_0^\pi [\sin^2\Theta-2\cos^2\Theta]^2\sin\Theta d\Theta. \tag{12}$$

A corresponding difference is added to S_U, namely,

$$\Delta S=\int_0^{x_+}\frac{\Delta}{x^2}dx=\frac{16}{35}\left(1-\frac{1}{(1+x_+)^7}\right)=0.29 \tag{13}$$

for $x_+=4/25$. This is to be compared with S_U which has the value 2.1. The correction is limited ($0<\Delta S<16/35$), depending on the value of x_+. Its inclusion would not change appreciably the results of our simpler approach.

For polarization broadening $F_1-F_{2p,0;2p,0}$ is replaced by $(1/2)F_{2s,0;2s,0}-F_{2p,0;2p,0}$. Since $\psi_{2s,0}$ like $\psi_{1s,0}$ is spherically symmetric, $(1/4)\Delta S$ is to be added to S_P of Sec. IV. Here the correction is not negligible and

$$\gamma_P/\gamma=0.061[0.081+(0.29/4)]=0.0094 \tag{14}$$

in place of (IV-32).

Our earlier treatment of quenching is rigorous. The form factor is

$$\begin{aligned}F_{1s,0;2p,0}&=\int e^{i\mathbf{K}\cdot\mathbf{r}}\psi_{1s,0}{}^{r'}\psi_{2p,0}{}^{r'*}d\mathbf{r}\\&=\int e^{i\mathbf{K}\cdot\mathbf{r}}\psi_{1s,0}{}^{r}\left[\cos\Theta\psi_{2p,0}{}^{r}+\frac{\sin\Theta}{\sqrt{2}}(e^{-i\Phi}\psi_{2p,+1}{}^{r}+e^{i\Phi}\psi_{2p,-1}{}^{r})\right]d\mathbf{r}\\&=\cos\Theta\int e^{i\mathbf{K}\cdot\mathbf{r}}\psi_{1s,0}{}^{r}\psi_{2p,0}{}^{r}d\mathbf{r}\\&=\cos\Theta F_{1s,0;2p,0}(\Theta=0),\end{aligned}$$

and the transition probability averaged over orientations is proportional to

$$\tfrac{1}{2}\int_0^\pi \cos^2\Theta\sin\Theta d\Theta[F_{1s,0;2p,0}(0)]^2=\tfrac{1}{3}[F_{1s,0;2p,0}(0)]^2.$$

This is the value used previously [V-(7)].

VII. MATRIX ELEMENTS WITH SPURIOUS DIVERGENCES

Some matrix elements occurring in the analysis of polarizing collisions appear to be imaginary and to diverge if the momentum transfer vector ($\mathbf{K}$) is zero. Hence a contradiction arises, for the integrals are obviously real. We shall see that they do not diverge if the perturbing electron charge is distributed over a finite volume.

This obviates the difficulty which prevented us from including all degenerate states in the polarization problem. Polarization by induction involves electron-enforced transitions between $\psi_\pm$ and $\psi_{2p,\pm1}$ (see Sec. IV). These contain a matrix element which we now single out for study. It is

$$C(\mathbf{K})_{2p,0;2s,0}=\frac{4\pi e^2}{K^2V}F_{2p,0;2s,0}, \tag{1}$$

where

$$F_{2p,0;2s,0}=\int e^{i\mathbf{K}\cdot\mathbf{r}}\psi_{2p0}{}^*\psi_{2s0}d\mathbf{r}.$$

Use has here been made of the formula

$$Z\equiv\int\frac{e^{i\mathbf{K}\cdot\mathbf{R}}}{|\mathbf{r}-\mathbf{R}|}d\mathbf{R}=\frac{4\pi}{K^2}e^{i\mathbf{K}\cdot\mathbf{r}}, \tag{2}$$

which was derived by [III-(9, 11)].

The form factor $F_{2p,0;2s,0}$ contains

$$\psi_{2s,0}=\frac{1}{4(2\pi)^{\frac{1}{2}}a^{\frac{3}{2}}}(2-r/a)e^{-r/2a}$$

and

$$\psi_{2p,0}=\frac{1}{4(2\pi)^{\frac{1}{2}}a^{5/2}}re^{-r/2a}\cos\theta.$$

In performing the integral over θ, we can see that the form factor is pure imaginary, as follows. Taking $\mathbf{K}$ along the polar axis of the atom and letting $\mu=\cos\theta$, we have

$$\int_{-1}^{+1}e^{iKr\mu}\mu d\mu=2i\left[\frac{\sin Kr}{(Kr)^2}-\frac{\cos Kr}{Kr}\right].$$

This yields

$$F_{2p,0;2s,0}=3iKa(x-1)/(1+x)^4; \tag{3}$$

and

$$C(K)_{2p,0;2s,0}=\frac{4\pi e^2}{VK^2}\left[\frac{3iKa(x-1)}{(1+x)^4}\right],\quad x=K^2a^2. \tag{4}$$

This element diverges as $1/K$ when K goes to zero, and it is imaginary. A look at the integral (1) for $\mathbf{K}=0$ tells us that the result should be real. We note that the form-factor alone does not introduce this difficulty, since it becomes zero. This suggests that the $1/K^2$ factor in (2) is the culprit.

We may view $C(\mathbf{K})_{n,n'}$ as the electrostatic interaction energy between atomic charge, $dq_a(\mathbf{r})=e\psi_n^*\psi_{n'}d\mathbf{r}$ and an electron charge density $dq_e=(e^{i\mathbf{K}\cdot\mathbf{R}}/V)ed\mathbf{R}$. The integral (2), when multiplied by e/V, then plays the role of the potential at $\mathbf{r}$ due to the charge density distribution $(e/V)e^{i\mathbf{K}\cdot\mathbf{R}}$. It satisfies Poisson's equation[20]:

$$\nabla^2\left(\frac{4\pi e}{VK^2}e^{i\mathbf{K}\cdot\mathbf{r}}\right)=-\frac{4\pi e}{V}e^{i\mathbf{K}\cdot\mathbf{r}}$$

For $K=0$, therefore, it is proper to evaluate $C_{nn'}$ by ordinary electrostatics, taking q_e to be a uniform spherical distribution extending to R_+. The mutual potential energy between q_e and $dq_a(\mathbf{r})$ is

$$dU=\frac{2\pi e}{V}(R_+^2-\tfrac{1}{3}r^2)dq_a(\mathbf{r}).$$

This leads to

$$C_{nn'}(K=0)=\int dU=\frac{2\pi e^2}{V}(R_+^2\delta_{nn'}-\tfrac{1}{3}\langle r^2\rangle_{nn'}). \quad (5)$$

V is the volume per free electron, R_+ some finite cut-off radius introduced by the over-all neutrality of the plasma. For $n=2p,0$ and $n'=2s,0$, the term $\langle r^2\rangle_{nn'}$ vanishes and therefore $C_{nn'}=0$.

To obtain $C_{nn'}$ for small K, we modify the calculation of formula (2), heeding the suggestion implied by the electrostatic consideration and using a finite cut-off radius. By the previous substitutions, (2) becomes

$$Z=\frac{4\pi}{K}e^{i\mathbf{K}\cdot\mathbf{r}}\int_0^{\rho_+}\sin K\rho d\rho,$$

where $\boldsymbol{\rho}=\mathbf{R}-\mathbf{r}$.

At large values $\rho\cong R$, which goes to R_+. We assume $KR_+\ll 1$; then

$$\int_0^{\rho_+}\sin K\rho d\rho\cong\int_0^{R_+}K\rho d\rho=KR_+^2/2.$$

Thus

$$Z\cong 2\pi e^{i\mathbf{K}\cdot\mathbf{r}}R_+^2, \quad KR_+\ll 1.$$

Consequently,

$$C_{n,n'}(KR_+\ll 1)=(2\pi R_+^2e^2/V)F_{n,n'} \quad (6)$$

and

$$C_{n,n'}(K=0)=(2\pi R_+^2e^2/V)\delta_{n,n'}, \quad (7)$$

which agrees with (5) with neglect of the second term, (the correction for the distributed charge). Evidently we have made an approximation in the transformation from $\mathbf{R}$ to $\boldsymbol{\rho}$ [in Eq. (2)], by which this term is lost. This was done by inexact treatment of the upper limit of integration, which depends on the angle θ_{rR}. If done correctly

$$Z\cong 2\pi e^{i\mathbf{K}\cdot\mathbf{r}}[R_+^2-\tfrac{1}{3}r^2], \quad (8)$$

provided $\mathbf{K}\cdot(\mathbf{R}_+-\mathbf{r})\ll 1$.

Thus the problem of the divergence as $K\to 0$ can be avoided by introduction of a plasma cutoff R_+.

VIII. POLARIZATION AND DEGENERATE LEVELS

Section V treated the small broadening effect resulting from polarization for a three state atom, i.e., polarization by reorientation. The matrix elements were selected deliberately to avoid the divergences associated with the degenerate levels of the preceding section. Now we are prepared to study collisions which couple the initial state ψ_+ to all atomic states.

Excitation to higher principal quantum numbers is not energetically possible for low electron energies such as that considered in our example [$\epsilon_i\approx(1/25)e^2/2a$]. Even where they are possible, they introduce no new features. On the other hand, quenching-transitions to the ground state do occur. As seen in Sec. V, the divergence poses no problem for them; the large amount of energy given to the electron keeps the minimum momentum transfer far enough from zero to avoid this difficulty. The results obtained in Sec. V for quenching are slightly modified when we regard the upper state, not as a random combination of degenerate states, but as the one corresponding to ψ_+. Our present concern with polarization requires this latter choice. The line width in question is

$$(\gamma_P)_{+;1s,0}/\gamma=\tfrac{1}{2}f\{(S_P)_{2s,0;1s,0}+\tfrac{1}{3}(S_P)_{2p,0;1s,0}\}$$
$$=\tfrac{1}{2}(0.061)\{0.032+0.021\}=0.0016, \quad (1)$$

provided $n=10^{14}$ and $\epsilon_i=(1/25)e^2/2a$. This should be compared with the previous

$$\gamma_Q/\gamma=0.0013. \quad (2)$$

Only for degenerate levels does the minimum momentum transfer approach zero, and we enter the region of spurious divergences. We meet this problem when considering the two undisplaced states, $\psi_{2p,\pm 1}$. We define these states in terms of functions with a polar axis z along the direction of the momentum transfer vector $\mathbf{K}$, as follows:

$$\psi_{2p,+1^{z'}}=-\frac{\sin\Theta}{\sqrt{2}}\psi_{2p,0^z}+\tfrac{1}{2}(\cos\Theta+1)e^{-i\Phi}\psi_{2p,+1^z}$$
$$+\tfrac{1}{2}(\cos\Theta-1)e^{i\Phi}\psi_{2p,-1^z},$$

and

$$\psi_{2p,-1^{z'}}=-\frac{\sin\Theta}{\sqrt{2}}\psi_{2p,0^z}+\tfrac{1}{2}(\cos\Theta-1)e^{-i\Phi}\psi_{2p,+1^z}$$
$$+\tfrac{1}{2}(\cos\Theta+1)e^{i\Phi}\psi_{2p,-1^z}.$$

[20] H. S. W. Massey and C. B. O. Mohr, Proc. Roy. Soc. (London) A132, 605 (1931).

The angles Θ and Φ determine the direction of z with respect to z'. Thus, along with

$$\psi_{+}{}^{z'}\equiv(1/\sqrt{2})(\psi_{2s,0}{}^{z'}+\psi_{2p,0}{}^{z'})$$

and $\psi_{-}{}^{z'}$, there are four states of principal quantum number $n=2$, which are normalized and orthogonal. We write similar expansions for $\psi_{2s,0}{}^{z'}$ and $\psi_{2p,0}{}^{z'}$:

$$\psi_{2s,0}{}^{z'}=\psi_{2s,0}{}^{z},$$

$$\psi_{2p,0}{}^{z'}=\cos\Theta\psi_{2p,0}{}^{z}+(1/\sqrt{2})\sin\Theta e^{-i\Phi}\psi_{2p,+1}{}^{z}+(1/\sqrt{2})\sin\Theta e^{i\Phi}\psi_{2p,-1}{}^{z}.$$

First, consider the form factor:

$$F_{2p,+1;+}(\Theta,\Phi)=\int e^{i\mathbf{K}\cdot\mathbf{r}}\psi_{2p,+1}{}^{z*}\psi_{+}{}^{z}d\mathbf{r}$$

$$=(1/\sqrt{2})[F_{2p,+1;2s,0}(\Theta,\Phi)+F_{2p,+1;2p,0}(\Theta,\Phi)],$$

with

$$F_{2p,+1;2s,0}(\Theta,\Phi)=-(\sin\Theta/\sqrt{2})F_{2p,0;2s,0}(0)$$

$$=-(\sin\Theta/\sqrt{2})[3iKa(x-1)/(1+x)^4]$$

and

$$F_{2p,+1;2p,0}(\Theta,\Phi)=-(1/\sqrt{2})\sin\Theta\cos\Theta F_{2p,0;2p,0}(0)$$

$$=(1/2\sqrt{2})(\cos\Theta+1)\sin\Theta F_{2p,+1;2p,+1}(0)+(1/2\sqrt{2})(\cos\Theta-1)\sin\Theta F_{2p,-1;2p,-1}(0)$$

$$=\frac{1}{\sqrt{2}}\sin\Theta\cos\Theta\frac{6x}{(1+x)^4}.$$

We note that if $\Theta=0$ this form factor vanishes. This tells us that collisions in which the momentum transfer is along the polar z' axis cannot change the atom from $\psi_{+}{}^{z'}$ to $\psi_{2p,\pm1}{}^{z'}$.

The transition probability is proportional to the absolute square of the form factor averaged over collision orientations:

$$\frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi}|F_{2p,+1;+}(\Theta,\Phi)|^2\sin\Theta d\Theta d\Phi=\tfrac{1}{6}\{|F_{2p,0;2s,0}(0)|^2+\tfrac{1}{5}[6x/(1+x)^4]^2\},$$

where $F(0)$ indicates that $\Theta=0$ and $\mathbf{K}$ is along the polar axis of the atom. As before, we can write

$$(\gamma_P)_{+;2p,+1}=2nv_i\sigma_i(S_P)_{+;2p,+1}, \tag{3}$$

with

$$(S_P)_{+;2p,+1}=\frac{1}{6}\int_{x_-}^{x_+}\frac{|F_{2p0,2s0}(0)|^2}{x^2}dx+\frac{6}{5}\int_0^{x_+}\frac{dx}{(1+x)^8}$$

$$=\tfrac{1}{6}(S_P)_{2p,0;2s,0}+\frac{6}{35}\left[1-\frac{1}{(1+x_+)^7}\right]. \tag{4}$$

The false divergence of $(S_P)_{2p,0;2s,0}$ at small x appears in this term.

It might be thought that it can be avoided because x_- is not zero but $a^2(k_i-k_f)^2$, and k_i cannot equal k_f in view of the shift which the + level has undergone relative to $2p,+1$; for conservation of energy requires $k_f{}^2-k_i{}^2=(2m/\hbar^2)\Delta E$. Inspection shows, however, that this is not the case for reasonable ion densities. Nor would the removal of a mathematical difficulty by the artifact of a static ion field be altogether satisfactory.

We therefore turn to the development of Sec. VII. From (VII-8),

$$\int\frac{e^{i\mathbf{K}\cdot\mathbf{R}}}{|\mathbf{R}-\mathbf{r}|}d\mathbf{R}\cong 2\pi e^{i\mathbf{K}\cdot\mathbf{r}}\left[R_+{}^2-\frac{r^2}{3}\right]. \tag{5}$$

If we neglect the $r^2\sim a^2$ contribution as being small compared to R_+ then [VII-(6)] gives

$$C_{2s,0;2p,0}=(4\pi e^2/V)(\tfrac{1}{2}R_+{}^2)F_{2s,0;2p,0}, \tag{6}$$

so that $R_+{}^2/2$ has taken the place of $1/K^2$, as is seen from (VII-1). This change prevents the divergence. We now know the matrix element at the two extremes of x. In these two regions,

$$|C_{2s,0;2p,0}|^2=(4\pi e^2a^2/V)^2|F_{2s,0;2p,0}|^2\times\begin{cases}\frac{1}{4}(R_+/a)^4 & \text{if } x(R_+/a)^2<1\\ 1/x^2 & \text{if } x(R_+/a)^2>1,\end{cases}$$

in view of VIII-(6) and VII-(1).

In order to average over orientations of $\mathbf{K}$ we use the previous expansions. Setting $\mu=\cos\Theta$,

$$\langle|F_{2s,0;2p,0}|^2\rangle_{Av}\equiv\tfrac{1}{2}\int_{-1}^{+1}|F_{2s,0;2p,0}(\Theta)|^2d\mu$$

$$=\tfrac{1}{2}\cdot\tfrac{2}{3}|F_{2s,0;2p,0}(0)|^2=\tfrac{1}{3}[9x(x-1)^2/(x+1)^8],$$

since $F_{2s,0;2p,0}$ is given by [VII-(3)]. Hence, when $x\ll1$, $\langle|F_{2s,0;2p,0}|^2\rangle_{Av}\cong3x$ and the matrix elements have the simple form:

$$\langle|C_{2s,0;2p,0}|^2\rangle_{Av}=(4\pi e^2a^2/V)^2\begin{cases}\frac{3}{4}x(R_+/a)^4 & \text{if } x(R_+/a)^2<1\\ 3/x & \text{if } x(R_+/a)^2>1.\end{cases}$$

Following the considerations of Sec. VII, it can be shown that $|C_{2s,0;2p,0}|^2$ is bounded in the middle region. Since

$$\left|\int_0^{\rho_+}\sin K\rho d\rho\right|\leqslant\left|\frac{2}{K}\right|,$$

we have

$$|C_{2s,0;2p,0}{}^{(0)}|^2\leqslant4\left|\frac{4\pi e^2}{VK^2}F_{2s,0;2p,0}(0)\right|^2\cong4\left(\frac{4\pi e^2a^2}{V}\right)^2\frac{9}{x}.$$

In Fig. 5, the curves for $|C|^2$ are extrapolated into the region of uncertainty. To avoid a detailed treatment of this intermediate region we use the extrapolated curves to their intersection. This intersection $\bar{x}$ occurs where $(R_+/a)^4(\bar{x}/4)=1/\bar{x}$; $\bar{x}\sim 2a^2n^{\frac{1}{2}}\sim 10^7$ for $n=10^{14}$ cm^{-3}. Consequently, neglecting x against 1, we find

$$(S_F)_{2p,0;2s,0}(0)\cong\tfrac{1}{4}(R_+/a)^4\int_0^{x}|F_{2s,0;2p,0}(0)|^2dx + \int_{\bar{x}}^{x_+}\left|\frac{F_{2s0,2p0}(0)}{x}\right|^2 dx$$

$$\cong\frac{9}{\bar{x}^2}\int_0^{\bar{x}}xdx+9\int_{\bar{x}}^{x_+}\frac{dx}{x}$$

$$\cong 9\left\{\frac{1}{2}+\ln\frac{x_+}{\bar{x}}\right\}\cong 9\left\{\frac{1}{2}+\ln\frac{0.16}{10^{-7}}\right\}$$

$$\cong 133, \qquad (7)$$

since $|F_{2s,0;2p,0}(0)|^2=9x(x-1)^2/(1+x)^8$. And finally, in view of Eq. (4) (with $x_+=0.16$),

$$(\gamma_P)_{+;2p,+1}/\gamma\sim f\{\tfrac{1}{6}\times 133+0.11\}=0.061\times 22\cong 1.3.$$

The main uncertainty in this number arises from the manner in which we obtained $\bar{x}$. We note, however, that an increase in $\bar{x}$ by a factor 100 decreases $(\gamma_P)_{+;2p,+1}/\gamma$ by only $\frac{3}{2}f\ln 100\cong 0.4$. This leaves the order of magnitude of the effect unchanged. In the other direction $\bar{x}$ cannot be pushed very far, since $\bar{x}(R_+/a)^2\sim 2$ has been used in (7) and the logarithmic divergence is limited by $\bar{x}(R_+/a)^2\sim 1$. R_+ may, of course, be larger than $n^{-1.3}$. An increase of R_+ by factor 100 decreases $\bar{x}$ by 10^{-4} but increases γ_P/γ by only 0.8.

We conclude that coupling between ψ_+ and $\psi_{2p,+1}$ gives rise to a polarization effect about 100 times as large as inclusion of the other states. This is because the coupling involves the matrix elements between $\psi_{2p,0}$ and $\psi_{2s,0}$. However, the effect does not diverge.

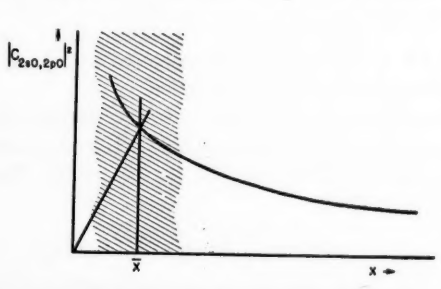


FIG. 5. A sketch of $|C_{2s0,2p0}|^2$, known at large and small $x=a^2K^2$ (K=momentum transfer). The matrix element in the shaded area is found by extrapolation from the extremities. The point of intersection is called $\bar{x}$.

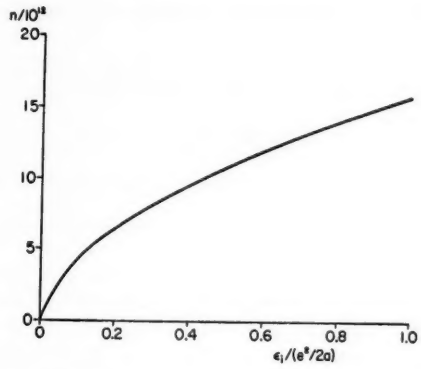


FIG. 6. Curve of density (n) *vs* electron energy (ϵ_i) delimiting the region of small n and large ϵ_i where electron broadening is negligible compared to the natural width (γ). The curve is the locus of points satisfying $(\gamma_P)_{2p0,2s0}=\gamma/10$.

In Fig. 6, we plot the locus of $\gamma_P/\gamma=1/10$ in order to indicate the region of electron densities and velocities in which their broadening may be neglected.

A word is in order about the line shifts, which are neglected in this treatment. The use of the resonance approximation, which ignores the variation of the matrix elements C with change in electron energy Ω, makes γ_P real. When this dependence is properly considered, an imaginary term is added to γ_P. This produces a line shift which will be discussed in a separate publication.

As a summary, we compare the several effects for the hydrogen $2p,0\rightarrow 1s,0$ transition, using $\gamma=3.12\times 10^8$ sec^{-1}, $n=10^{14}$ cm^{-3}, $\epsilon_i=(1/25)e^2/2a$:

III-(38)	Universal broadening	$\gamma_U/\gamma=$ 0.13
VIII-(16)	Polarization by re-orientation (involving the matrix elements $2p0\rightarrow 2p\pm 1$)	$\gamma_P/\gamma=$ 0.027
V-A-(7)	Quenching ($2p0\rightarrow 1s0$)	$\gamma_Q/\gamma=$ 0.0013
VIII-(16)	Polarization by induction (involving the large element $2p0\rightarrow 2s0$)	$\gamma_P/\gamma=$ 2.7
V-B-(8)	Stark broadening by ions	$\gamma_{\text{Stark}}/\gamma\cong 80$

Thus for the Lyman α line, where there is a linear Stark effect, the ion broadening is the dominant effect, as expected. But in other instances, to be treated in further publications, the electrons can make important contributions.

APPENDIX

The principal symbols used in this paper are as follows: E=energy of atomic state; ϵ=energy of electron: ϵ_i=initial energy, ϵ_μ=final energy; $\mathbf{k}$=wave-number vector for electron; $\mathbf{K}=\mathbf{k}_\mu-\mathbf{k}_i$=momentum-transfer vector; v=velocity of electron; $a=\hbar^2/me^2$;

$x=a^2K^2$; $x_+=a^2(k_\mu+k_i)^2$; $x_-=a^2(k_\mu-k_i)^2$; n=number density of electrons.

Throughout the paper, line widths resulting from different causes are written in the Lorentz form:

$$\gamma=2nv\sigma S,$$

where σ is a cross section and S a numerical efficiency factor. In particular,

$$S_U\equiv\int_0^{x+}\frac{|F_{11}-F_{22}|^2}{x^2}dx,\quad S_P\equiv\int_0^{x+}\frac{|F_{nm}|^2}{x^2}dx,$$

and

$$F_{nm}=\int e^{i\mathbf{K}\cdot\mathbf{r}}\psi_n^*(r)\psi_m(r)d\mathbf{r}.$$

PHYSICAL REVIEW VOLUME 98, NUMBER 2 APRIL 15, 1955

Extension of Hund's Rule*

G. F. KOSTER†
Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received September 20, 1954)

In this note we show that in an n-electron system, if we assume that these n electrons occupy n orthogonal one-electron orbitals, the state of highest multiplicity has the lowest energy.

ONE of the interesting results that comes out of the calculation of the energy levels of atoms, molecules, and crystals is the total spin (multiplicity) of the ground state. In general the question of the multiplicity of the ground state cannot be answered without lengthy calculations. It is, however, sometimes possible to answer this question without any calculation at all (at least to within a certain approximation). We know of Hund's rule of atomic spectra. Part of this rule states that of all the levels arising from a given spatial configuration, the state of highest multiplicity lies lowest. A rule similar to this can be proved quite generally.[1]

We imagine that we have n distinct one-electron spatial orbitals $u_1, u_2, \cdots u_n$. In each of these we put one electron with either spin up or spin down. (The presence of other orbitals which are doubly filled does not influence the validity of the arguments given below and will be dropped from consideration.) We also assume that the orbitals under consideration are orthogonal.

$$\int u_i^*(\mathbf{r})u_j(\mathbf{r})d\tau=\delta_{ij}. \tag{1}$$

For our n-electron problem, by taking products of these one-electron states with spin up or down, we can form 2^n product functions which we use as the basis of approximation for the wave function for our n electron system. We must now form the correct linear combinations of these product functions to form states which are antisymmetric with respect to permutations of all of the electron coordinates and which have definite multiplicities. The multiplicities will range from singlets (n even) to states of multiplicity $n+1$. We shall lose no generality by restricting ourselves to all those states with the same z component of total spin. Of these only one corresponds to the state of highest multiplicity, whereas for any other multiplicity there is, in general, more than one state of this multiplicity. We shall now show that there is no state that lies lower in energy than the state of highest multiplicity. The problem which confronts us now is to diagonalize the matrix of our Hamiltonian between states of a given multiplicity in the basis of antisymmetric functions. (There is, of course, no interaction between the states of different multiplicities.) We assume that our Hamiltonian in this case is spin-free and is given by

$$H=\sum_1^n(i)[-\nabla_i^2+f(i)]+\tfrac{1}{2}\sum_1^n(i\neq j)g_{ij}. \tag{2}$$

Here $f(i)$ is the one-electron potential and g_{ij} is the interelectronic Coulomb interaction.

The problem as stated above is in just the form which can easily be set up in terms of the Dirac vector model.[2] For a given multiplicity the matrix of the Hamiltonian is given by

$$\mathcal{H}_{\lambda\sigma}{}^S=\text{const}+\tfrac{1}{2}\sum(i\neq j)[-J_{ij}(P_{ij}{}^S)_{\lambda\sigma}]. \tag{3}$$

* The research in this paper was supported jointly by the Army, Navy, and Air Force under contract with the Massachusetts Institute of Technology.

† Staff Member, Lincoln Laboratory, Massachusetts Institute of Technology.

[1] This extension of Hund's rule has been published in the Quarterly Progress Report of the Solid-State and Molecular Theory Group at Massachusetts Institute of Technology, July 15, 1953, p. 37. P.-O. Löwdin [Phys. Rev. **97**, 1474, 1490, 1509 (1955)] has also presented a proof by a different method.

[2] P. A. M. Dirac, Proc. Roy. Soc. (London) **A123**, 714 (1929). For further references and a complete discussion of this method see F. M. Corson, *Perturbation Methods in the Quantum Mechanics of n-Electron Systems* (Hafner Publishing Company, New York, 1950).

Here $\mathcal{H}_{\lambda\sigma}{}^{S}$ is the λ, σ matrix element in the Hamiltonian matrix which connects states of total spin S. $(P_{ij}{}^{S})$ are unitary matrices which represent the permutation of the spin coordinates of electron i with those of electron j. These matrices are part of an irreducible representation of the permutation group and the superscript S which denotes the total spin also specifies the particular irreducible representation. The order of these matrices depends on the multiplicity. They are of the same order as the number of states of a given multiplicity which can arise from n singly occupied orbitals. The order of these matrices are $n!/[(\frac{1}{2}n+S)!(\frac{1}{2}n-S)!] - n!/[(\frac{1}{2}n+S+1)!(\frac{1}{2}n-S-1)!]$. The matrices have only one row and column for the states of highest multiplicity ($S=\frac{1}{2}n$). The constant in the expression (3) contains all the one-electron interactions as well as the interaction of the orbitals which are unpaired $(u_1,u_2,\cdots u_n)$ with the orbitals which are paired and the interactions of the paired orbitals amongst themselves. This constant is independent of the multiplicity of the state under consideration. J_{ij} is the customary exchange interaction,

$$J_{ij}=\int u_i^*(\mathbf{r}_1)u_j(\mathbf{r}_1)g_{12}u_i(\mathbf{r}_2)u_j^*(\mathbf{r}_2)d\tau_1 d\tau_2. \quad (4)$$

This quantity is greater than zero since it represents the electrostatic interaction of a charge distribution with itself.

First let us consider the state of highest multiplicity. As we have mentioned the matrix of the Hamiltonian has only one row and column. The matrices representing the permutations P_{ij} have as their component unity. Therefore the energy of the state of highest multiplicity is [apart from the constant in (3) which is independent of multiplicity]:

$$\mathcal{H}_{11}{}^{\frac{1}{2}n}=-\tfrac{1}{2}\sum(i\neq j)J_{ij}. \quad (5)$$

For any other multiplicity, as we have seen, the matrices $(P_{ij}{}^{S})_{\lambda\sigma}$ are of larger order. Let us consider the diagonal energies of all the states of a given multiplicity. These energies are the diagonal elements of the matrix in Eq. (3):

$$\mathcal{H}_{\sigma\sigma}{}^{S}=-\tfrac{1}{2}\sum(i\neq j)J_{ij}(P_{ij}{}^{S})_{\sigma\sigma}. \quad (6)$$

Using the unitary nature of the matrices $(P_{ij}{}^{S})_{\lambda\sigma}$ and the fact that $J_{ij}=J_{ji}$, Eq. (6) can be rewritten in the form

$$\mathcal{H}_{\sigma\sigma}{}^{S}=-\tfrac{1}{2}\sum(i\neq j)J_{ij}[\tfrac{1}{2}(P_{ij}{}^{S})_{\sigma\sigma}+\tfrac{1}{2}(P_{ij}{}^{S})_{\sigma\sigma}{}^{*}]. \quad (7)$$

In the above, we have also made use of the fact that the inverse of the permutation P_{ij} is the permutation P_{ji}. Once again using the unitary nature of the matrices representing the permutations, we know that

$$\tfrac{1}{2}[(P_{ij}{}^{S})_{\sigma\sigma}+(P_{ij}{}^{S})_{\sigma\sigma}{}^{*}]\leqslant 1. \quad (8)$$

It is clear, therefore, from this fact and the fact that $J_{ij}>0$ that

$$\mathcal{H}_{11}{}^{\frac{1}{2}n}=-\tfrac{1}{2}\sum(i\neq j)J_{ij}<\mathcal{H}_{\sigma\sigma}{}^{S}$$
$$=-\tfrac{1}{2}\sum(i\neq j)J_{ij}[\tfrac{1}{2}(P_{ij}{}^{S})_{\sigma\sigma}+\tfrac{1}{2}(P_{ij}{}^{S})_{\sigma\sigma}{}^{*}]. \quad (9)$$

We have therefore shown that there is no state lower than the state of highest multiplicity under the assumptions of this proof.

One might ask if the process of diagonalizing the Hamiltonian matrix (3) will push some state lower than the state of highest multiplicity. This is not the case since in diagonalizing the matrix (3) we use a unitary transformation which does not destroy the unitary property of $(P_{ij}{}^{S})_{\lambda\sigma}$. This unitary property was all that was necessary to prove the theorem. P.-O. Löwdin[1] has, by a different method, presented a proof which is less complete but which gives the explicit energy values for a limited form of configurational interaction.

One might well ask how we could upset the result of this theorem to get a state of some other multiplicity to lie lower. In order to do this, we must include in our approximate wave function other states which consist of having some of our n orbitals doubly occupied and others empty.

Electric Quadrupole Interaction and Spin Echoes in Crystals

T. P. Das and A. K. Saha
Institute of Nuclear Physics, Calcutta, India

(Received September 27, 1954; revised manuscript received November 30, 1954)

Bloom and Norberg and also Hahn and Herzog have observed spin echoes due to electric quadrupole interaction alone and also in the presence of a weak magnetic field in addition to electric quadrupole interaction, with Cl^{35} (spin $I=3/2$) nuclei in $NaClO_3$ crystal. They have explained these echoes theoretically for the particular case of $I=3/2$. We have calculated the spin-echo amplitudes to be expected under the above experimental conditions for nuclei of any general spin in crystals with axial symmetry. The opposite case of spin echoes in a strong magnetic field in the presence of weak electric quadrupole interaction is also discussed. The theoretical investigation leads to "slow beats" in this case also.

INTRODUCTION

In the conventional spin-echo experiment,[1] we have a strong steady magnetic field, perpendicular to which a radio-frequency magnetic field is applied, and the free induction and echo signals following the application of the radio-frequency field in short pulses, is studied. However, in the case of many crystals, some of the nuclei have considerable electric quadrupole interaction with the surroundings. In such cases one can think of spin-echo experiments where the electric field gradient inside the crystal takes the place of the steady magnetic field. Such experiments have recently been carried out successfully by Hahn and Herzog[2] and also by Bloom and Norberg[3,4] in the case of chlorine nuclei in $NaClO_3$ crystal. They have also observed the modulation pattern obtained when a steady weak magnetic field is applied at an angle θ_0 to the axis of symmetry of the crystal. They have explained their observed results theoretically using a method similar to the one used earlier by Hahn and Maxwell.[5] We shall show in this paper how their results can be explained and extended to the case of nuclei with any general spin I, using the density matrix method of calculating spin-echo amplitudes developed by us in an earlier paper.[6] This paper is divided into three sections. In Sec. I, we take up the case of spin-echo signals for pure quadrupole coupling. In Sec. II, we examine the effect of a weak magnetic field superimposed on the electric quadrupole coupling, at an angle θ_0 to the axis of symmetry of the crystal. In Sec. III, the opposite case is treated, *viz.*, the one in which the electric quadrupole effect is small but the applied magnetic field is strong. As in our previous paper,[6] which we shall henceforth refer to as Paper I, we have not included the effects of the relaxation times T_1 and T_2, because from a phenomenological point of view, we expect their effects on the spin echoes in the present case to be the same as on the conventional spin echoes.[1] In a later paper we shall try to examine how the concept of density matrix[7] may be used in the context of the Bloembergen picture[8,9] of relaxation forces to introduce their effects directly into the spin-echo signal amplitudes.

SECTION I. PURE QUADRUPOLE CASE

As explained in Paper I, our method of calculating the spin-echo amplitudes essentially involves solution of the time-dependent Schrödinger equations for the spin system in the presence and absence of the rf pulses. If the Hamiltonian due to the electric quadrupole interaction be $\mathcal{H}_0$ and the additional perturbation due to the applied radio-frequency field be $\mathcal{H}'$, then if R and D represent the transformation matrices in the presence and absence of pulses respectively, then the Schrödinger equations in the two cases will be

$$\begin{aligned} i\hbar\frac{dR}{dt}&=(\mathcal{H}_0+\mathcal{H}')R,\\ i\hbar\frac{dD}{dt}&=\mathcal{H}_0 D. \end{aligned} \tag{1}$$

We shall refer to R in the future as the "transition matrix" depicting the transformation in the state of the spin-system in the presence of a rf pulse and to D as the "free-precession matrix" depicting the transformation in the state of the spin-system in the absence of the rf pulses. If we now apply n rf pulses, then the net transformation matrix S, termed the "echo matrix," will be given by

$$S=D_nR_nD_{n-1}R_{n-1}\cdots D_2R_2D_1R_1, \tag{2}$$

the suffixes referring to the successive pulses. The density matrix[6] $\rho(t)$ for the spin system after the passage of the pulses will then be related to the initial density matrix before the pulses are applied by the

[1] E. L. Hahn, Phys. Rev. **80**, 580 (1950).
[2] E. L. Hahn and B. Herzog, Phys. Rev. **93**, 639 (1954).
[3] M. Bloom and R. Norberg, Phys. Rev. **93**, 638 (1954).
[4] M. Bloom, Phys. Rev. **94**, 1396 (1954).
[5] E. L. Hahn and D. E. Maxwell, Phys. Rev. **88**, 1070 (1952).
[6] Das, Saha, and Roy, Proc. Roy. Soc. (London) **A227**, 407 (1955).
[7] P. A. M. Dirac, *Quantum Mechanics* (Clarendon Press, Oxford, 1947), third edition, p. 132.
[8] Bloembergen, Purcell, and Pound, Phys. Rev. **73**, 679 (1948).
[9] R. K. Wangsness and F. Bloch, Phys. Rev. **89**, 728 (1953), have introduced T_1 and T_2 in the differential equations for the density matrix under certain restricted conditions and without using the detailed Bloembergen picture.

relation

$$\rho(t)=S\rho(0)S^{-1}. \tag{3}$$

In most experiments, where the rf coil producing the rf field is also the detector, we need the expectation value of I_x, the x component of the spin of the members of the spin system, after the passage of pulses. This is given by

$$\langle I_x\rangle=\mathrm{Tr}\{\rho(t)I_x\}=\mathrm{Tr}\{S\rho(0)S^{-1}I_x\}, \tag{4}$$

The Hamiltonian in the case of pure quadrupole coupling may be written down, in the irreducible tensor scalar product form[10,11] as

$$\mathfrak{H}_0=\sum_r (Q)_r^2(\nabla\mathbf{E})_{-r}^2, \tag{5}$$

where the components $(Q)_r^2$ of the electric quadrupole moment tensor of the nucleus and $(\nabla\mathbf{E})_r^2$ of the field-gradient tensor are given by

$$\begin{aligned}(Q)_0^2&=\frac{eQ}{2I(2I-1)}(3I_0^2-\mathbf{I}^2),\\ (Q)_{\pm1}^2&=\mp(\tfrac{1}{2}\sqrt{6})\frac{eQ}{2I(2I-1)}\{I_\pm,I_0\},\\ (Q)_{\pm2}^2&=(\tfrac{1}{2}\sqrt{6})\frac{eQ}{2I(2I-1)}(I_\pm)^2,\end{aligned} \tag{6}$$

and

$$\begin{aligned}(\nabla\mathbf{E})_0^2&=\tfrac{1}{4}eq(3\cos^2\theta_0-1),\\ (\nabla\mathbf{E})_{\pm1}^2&=\pm(\tfrac{1}{4}\sqrt{6})eq\sin\theta_0\cos\theta_0\exp(\pm i\phi_0),\\ (\nabla\mathbf{E})_{\pm2}^2&=(\tfrac{1}{8}\sqrt{6})eq\sin^2\theta_0\exp(\pm2i\phi_0),\end{aligned} \tag{7}$$

with $I_0=I_z$ and $I_\pm=I_x\pm iI_y$; the notation $\{A,B\}$ refers to the anti-commutator $(AB+BA)$. Q is the scalar quadrupole moment defined by Casimir,[12] depending on the charge distribution within the nucleus. q is similarly the scalar quantity[10] defining the components of the field gradient tensor for axial symmetry depending on the charge distribution around the nucleus. ϕ_0 and θ_0 refer respectively to the azimuth and colatitude of the axis of the symmetry in the coordinate system chosen. If we now choose our coordinate system with the axis of symmetry as the z axis, we then have $\theta_0=0$, and we get,[13]

$$\mathfrak{H}_0=P\hbar(3I_0^2-\mathbf{I}^2), \tag{8}$$

where

$$P=\frac{e^2Qq}{4\hbar I(2I-1)}. \tag{9}$$

Evidently the energy levels will be specified by the eigenvalues of I_z and the levels corresponding to eigenvalues $+m$ and $-m$ will coincide. We shall thus have in all, $(I+1)$ distinct energy levels for integral I and $(I+\frac{1}{2})$ for half-integral I. The spacing between the adjacent levels will be unequal, and we shall therefore have in all, I resonance frequencies for integral I, and $(I-\frac{1}{2})$ for half-integral, corresponding to the energy differences between successive levels [see Figs. 1 (A and C)]. The resonance frequency corresponding to the transitions $m\rightarrow(m-1)$ and $-m\rightarrow-(m-1)$ will thus be given by

$$\omega(m)=3P[m^2-(m-1)^2]=3P(2m-1). \tag{10}$$

Let the applied radio-frequency be that corresponding to the transitions,

$$\pm m_1\rightarrow\pm(m_1-1).$$

For brevity, we introduce the following notations:

$$\begin{aligned}E(m_1)&=aP\hbar,\\ E(m_1-1)&=bP\hbar,\\ a&=3m_1^2-I(I+1),\\ b&=3(m_1-1)^2-I(I+1),\end{aligned} \tag{11}$$

so that the resonance frequency in question is

$$\omega(m_1)=P(a-b). \tag{12}$$

If the radio-frequency magnetic field be applied at right angles to the symmetry axis, (i.e., the z-axis), then calling its direction that of x, the perturbing Hamiltonian will be

$$\mathfrak{H}_1(t)=-\gamma\hbar I_x\cdot H_r\cos\omega(t-t_0), \tag{13}$$

where I_x is the x-component of the nuclear spin and $2H_r$ the amplitude of the radio-frequency field, ω being its frequency. Now, following the usual method[14] for dealing with time-dependent Hamiltonians, we have the matrix R given by,

$$\begin{aligned}R(t,t_0)&=\exp[-(i/\hbar)\mathfrak{H}_0(t-t_0)]\cdot R^*(t,t_0),\\ R^*(t,t_0)&=\exp\left[-(i/\hbar)\int_{t_0}^{t}\mathfrak{H}_1^*(t')dt'\right],\\ \mathfrak{H}_1^*(t')&=\exp[(i/\hbar)\mathfrak{H}_0(t'-t_0)]\cdot\mathfrak{H}_1(t')\\ &\quad\times\exp[-(i/\hbar)\mathfrak{H}_0(t-t_0)].\end{aligned} \tag{14}$$

In the present case, using (13),

$$\begin{aligned}\mathfrak{H}_1^*(t')=-\tfrac{1}{2}\omega_1\hbar\{&\exp(3iP(t'-t_0)(2I_0-1))\\ &\cdot I_+[\exp(i\omega(t'-t_0))+\exp(-i\omega(t'-t_0))]\\ &+\exp(-3iP(t'-t_0)(2I_0+1))\cdot I_-[\exp(i\omega(t'-t_0))\\ &+\exp(-i\omega(t'-t_0))]\}.\end{aligned} \tag{15}$$

Using $\omega=\omega(m_1)$ and putting $t-t_0=t_w$, the rf pulse

[10] R. V. Pound, Phys. Rev. **79**, 685 (1950).
[11] Refer to M. K. Banerjee and A. K. Saha, Proc. Roy. Soc. (London) **A224**, 472 (1954). We have followed their notation closely.
[12] H. B. G. Casimir, Arch. du Musée Teyler, **8**, 202 (1936).
[13] R. Bersohn, J. Chem. Phys. **20**, 1505 (1952).
[14] P. A. M. Dirac, *Quantum Mechanics* (Clarendon Press, Oxford, 1947), third edition, p. 173.

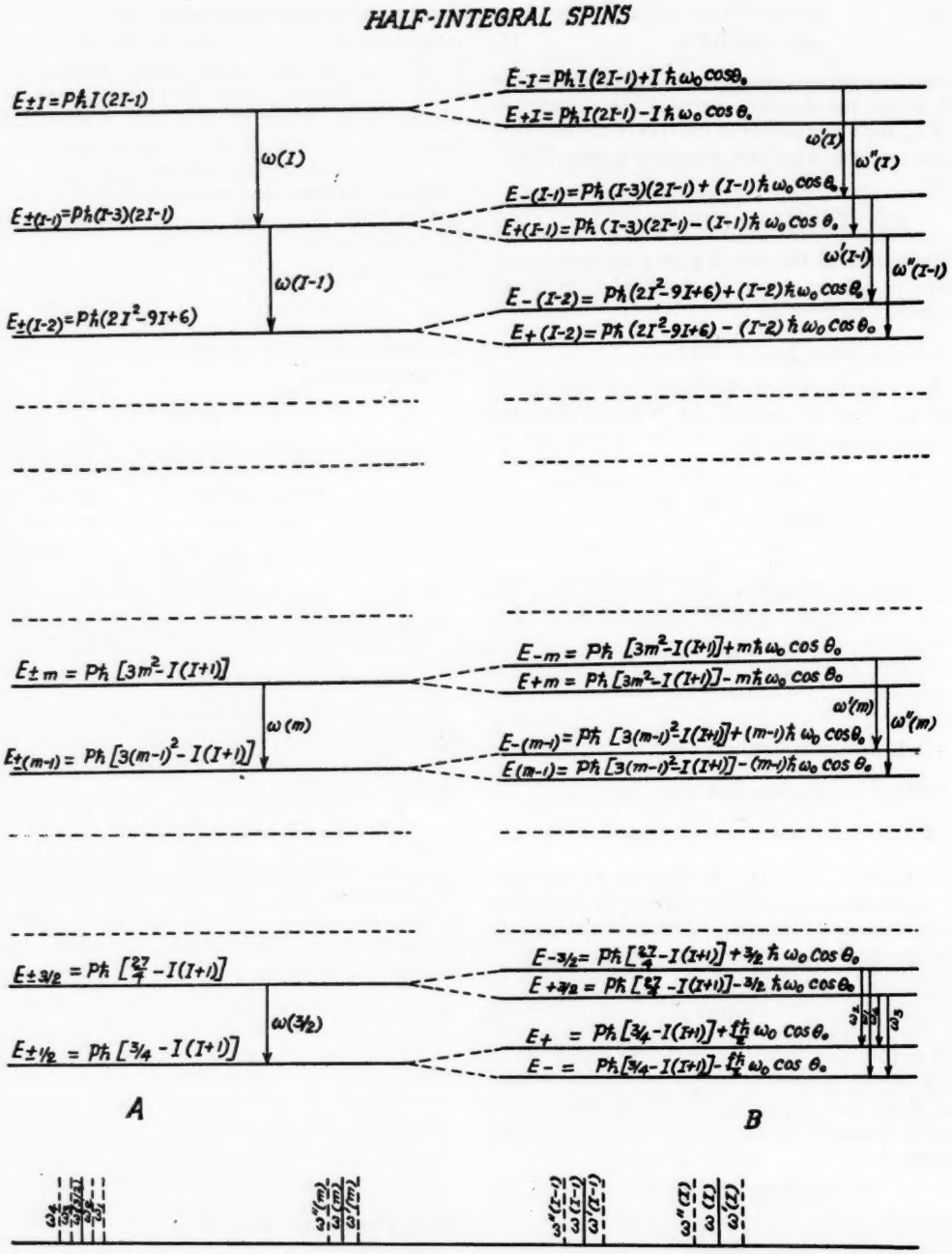

HALF-INTEGRAL SPINS
$E_{\pm I} = P\hbar I(2I-1)$
$E_{-I} = P\hbar I(2I-1) + I\hbar\omega_0\cos\theta_0$
$E_{+I} = P\hbar I(2I-1) - I\hbar\omega_0\cos\theta_0$
$\omega(I)$
$\omega'(I)$
$\omega''(I)$
$E_{\pm(I-1)} = P\hbar(I-3)(2I-1)$
$E_{-(I-1)} = P\hbar(I-3)(2I-1) + (I-1)\hbar\omega_0\cos\theta_0$
$E_{+(I-1)} = P\hbar(I-3)(2I-1) - (I-1)\hbar\omega_0\cos\theta_0$
$\omega(I-1)$
$\omega'(I-1)$
$\omega''(I-1)$
$E_{\pm(I-2)} = P\hbar(2I^2-9I+6)$
$E_{-(I-2)} = P\hbar(2I^2-9I+6) + (I-2)\hbar\omega_0\cos\theta_0$
$E_{+(I-2)} = P\hbar(2I^2-9I+6) - (I-2)\hbar\omega_0\cos\theta_0$
$E_{\pm m} = P\hbar[3m^2 - I(I+1)]$
$E_{-m} = P\hbar[3m^2 - I(I+1)] + m\hbar\omega_0\cos\theta_0$
$E_{+m} = P\hbar[3m^2 - I(I+1)] - m\hbar\omega_0\cos\theta_0$
$\omega(m)$
$\omega'(m)$
$\omega''(m)$
$E_{\pm(m-1)} = P\hbar[3(m-1)^2 - I(I+1)]$
$E_{-(m-1)} = P\hbar[3(m-1)^2 - I(I+1)] + (m-1)\hbar\omega_0\cos\theta_0$
$E_{(m-1)} = P\hbar[3(m-1)^2 - I(I+1)] - (m-1)\hbar\omega_0\cos\theta_0$
$E_{\pm 3/2} = P\hbar[\frac{27}{4} - I(I+1)]$
$E_{-3/2} = P\hbar[\frac{27}{4} - I(I+1)] + \frac{3}{2}\hbar\omega_0\cos\theta_0$
$E_{+3/2} = P\hbar[\frac{27}{4} - I(I+1)] - \frac{3}{2}\hbar\omega_0\cos\theta_0$
$\omega(3/2)$
$E_{\pm 1/2} = P\hbar[\frac{3}{4} - I(I+1)]$
$E_{+} = P\hbar[\frac{3}{4} - I(I+1)] + \frac{1}{2}\hbar\omega_0\cos\theta_0$
$E_{-} = P\hbar[\frac{3}{4} - I(I+1)] - \frac{1}{2}\hbar\omega_0\cos\theta_0$
A
B

FIG. 1. A and B are energy-level diagrams for half-integral spins resulting from pure electric quadrupole interaction only, and also in the presence of a weak steady magnetic field. The solid lines in the lower part of the figure represent the pure quadrupole resonance frequencies and the dotted lines the resonance lines in the presence of the weak magnetic field.

width, the matrix elements of R may be easily found from (14) and (15). Thus we have

$$\langle m|R|m'\rangle = \exp\{-iPl_w[3m^2 - I(I+1)]\}\delta_{mm'}, \quad \text{for } m \neq \pm m_1, \pm(m_1-1),$$

$$\langle \pm m_1|R|m'\rangle = \cos(x\xi/2)\exp(-iPal_w)\delta_{m',\pm m_1} + i\sin(x\xi/2)\exp(-iPal_w)\delta_{m',\pm(m_1-1)}, \tag{16}$$

$$\langle \pm(m_1-1)|R|m'\rangle = \cos(x\xi/2)\exp(-iPbl_w)\delta_{m',\pm(m_1-1)} + i\sin(x\xi/2)\exp(-iPbl_w)\delta_{m',\pm m_1},$$

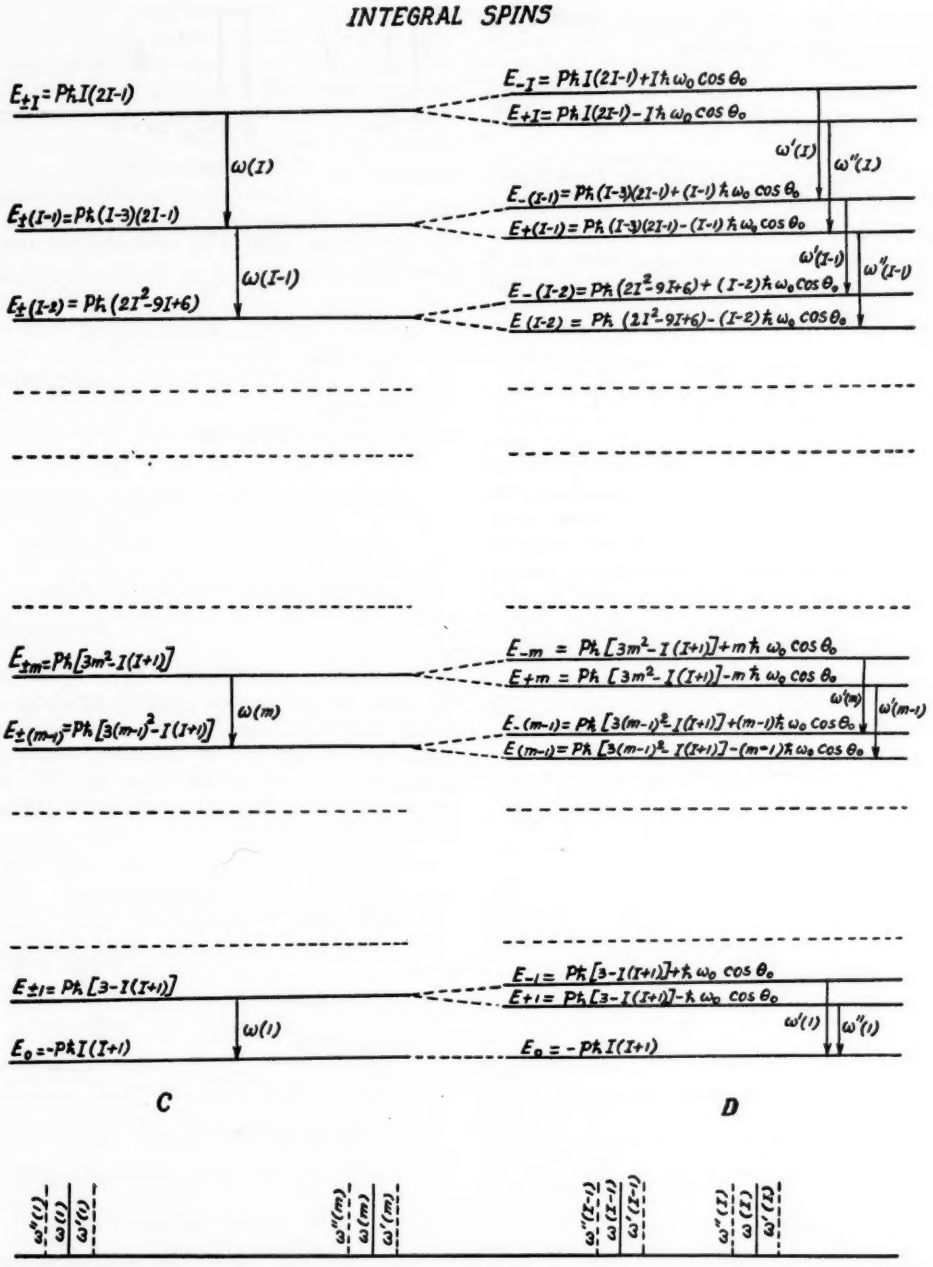


FIG. 1 (*continued*). C and D are energy-level diagrams for integral spins resulting from pure electric quadrupole interaction only, and also in the presence of a weak steady magnetic field. The solid lines in the lower part of the figure represent the pure quadrupole resonance frequencies and the dotted lines the resonance lines in the presence of the weak magnetic field.

where

$$x=[(I-m_1)(I+m_1+1)]^{\frac{1}{2}}. \tag{17}$$

In the absence of the rf field pulse, the free precession is given by the free precession matrix D, *viz.*,

$$D(t,t_0)=\exp[-(i/\hbar)\mathcal{H}_0(t-t_0)],$$

so that

$$\langle m|D(t,t_0)|m'\rangle = \exp\{-iP(t-t_0)[3m^2-I(I+1)]\}\delta_{mm'}, \tag{18}$$

for all m.

The density matrix at start, $\rho(0)$, is given by a normalized Boltzmann distribution,

$$\rho(0)=\frac{N}{2I+1}\exp\left[-\frac{P\hbar}{k\Theta}(3I_z^2-\mathbf{I}^2)\right], \tag{19}$$

where Θ is the absolute temperature and k the Boltzmann constant, N being the number of nuclei in the sample. If the spin-echo matrix at the end of a number of pulses be S, then we have the expectation value of I_x after these pulses given by (3). Remembering that $P\hbar/k\Theta\ll 1$, and S and I_x both commute with $\mathbf{I}^2$, we shall have

$$\langle I_x\rangle=\frac{N}{2I+1}\mathrm{Tr}\left(-\frac{3P\hbar}{k\Theta}SI_z^2S^{-1}I_x\right). \tag{20}$$

Using a pulse pattern of the type shown in Fig. 2 with pulse width t_w satisfying the usual conditions,[1] we have, following the procedure outlined in Paper I, the master matrices at the end of one and two pulses respectively given by the following equations.

Single pulse: $\tau<t<t_w$ (free induction signal)

$$\begin{aligned}&\langle m|S_{\mathrm{I}}|m'\rangle=\exp\{-iPt[3m^2-I(I+1)]\}\delta_{mm'},\\ &\qquad\text{for}\quad m\neq\pm m_1,\pm(m_1-1),\\ &\langle\pm m_1|S_{\mathrm{I}}|m'\rangle=a_1\delta_{m',\pm m_1}+c_1\delta_{m',\pm(m_1-1)},\\ &\langle\pm(m_1-1)|S_{\mathrm{I}}|m'\rangle=d_1\delta_{m',\pm m_1}+b_1\delta_{m',\pm(m_1-1)},\end{aligned} \tag{21}$$

with

$$\begin{aligned}a_1&=\exp(-iPat)\cdot\cos(x\xi/2),\\ b_1&=\exp(-iPbt)\cdot\cos(x\xi/2),\\ c_1&=i\exp(-iPat)\cdot\sin(x\xi/2),\\ d_1&=i\exp(-iPbt)\cdot\sin(x\xi/2).\end{aligned} \tag{22}$$

Two pulses: $t=(\kappa+\tau)$ (spin-echo signal)

$$\begin{aligned}&\langle m|S_{\mathrm{II}}|m'\rangle=\exp\{-iPt[3m^2-I(I+1)]\}\delta_{mm'},\\ &\qquad\text{for}\quad m\neq\pm m_1,\pm(m_1-1),\\ &\langle\pm m_1|S_{\mathrm{II}}|m'\rangle=a_2\delta_{m',\pm m_1}+c_2\delta_{m',\pm(m_1-1)},\\ &\langle\pm(m_1-1)|S_{\mathrm{II}}|m'\rangle=b_2\delta_{m',\pm m_1}+d_2\delta_{m',\pm(m_1-1)},\end{aligned} \tag{23}$$

where

$$\begin{aligned}a_2&=\cos^2(x\xi/2)\exp[-iPa(\kappa+\tau)]\\ &\quad-\sin^2(x\xi/2)\exp[-iP(a\kappa+b\tau)],\\ b_2&=\tfrac{1}{2}i\sin(x\xi)\{\exp[-iPb(\kappa+\tau)]\\ &\quad+\exp(-iP(b\kappa+a\tau)]\},\\ c_2&=\tfrac{1}{2}i\sin(x\xi)\{\exp[-iPa(\kappa+\tau)]\\ &\quad+\exp(-iP(a\kappa+b\tau)]\},\\ d_2&=\cos^2(x\xi/2)\exp[-iPb(\kappa+\tau)]\\ &\quad-\sin^2(x\xi/2)\exp[-iP(b\kappa+a\tau)].\end{aligned} \tag{24}$$

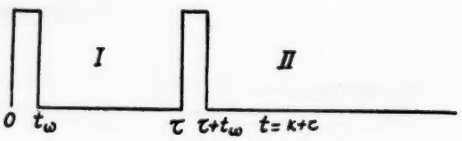


FIG. 2. Applied rf pulses.

Using Eqs. (21) and (23) together with (20), we get the expectation value of I_x after one and two pulses respectively as

$$\langle I_x\rangle=\frac{3NP\hbar}{(2I+1)k\Theta}x(2m_1-1)\sin(\omega(m_1)\cdot t),$$

(single pulse) (25)

$$\begin{aligned}\langle I_x\rangle=&\frac{3NP\hbar}{(2I+1)k\Theta}x(2m_1-1)\\ &\times\{\sin(x\xi)\cos^2(x\xi/2)\sin[\omega(m_1)\cdot t]\\ &-\sin(x\xi)\sin^2(x\xi/2)\sin[\omega(m_1)\cdot(t-2\tau)]\\ &+\sin(x\xi)\cos(x\xi)\sin[\omega(m_1)\cdot(t-\tau)]\},\end{aligned}$$

(two pulses). (26)

In case there is an inhomogeneity in q, and hence in P, there will be a corresponding inhomogeneity in $\omega(m_1)$. As in the case of echoes in a strong magnetic field, we can assume a Gaussian distribution in $\Delta\omega=\omega(m_1)-\omega$, with an rms value of $1/T_2^*$; we then get, on integrating the right-hand sides of (25) and (26) over $\Delta\omega$,

$$\langle I_x\rangle=\frac{3NP\hbar}{(2I+1)k\Theta}x(2m_1-1)\sin(x\xi)\sin\omega t\times\exp(-t^2/2T_2^{*2}),$$

(single pulse), (27)

$$\begin{aligned}\langle I_x\rangle=&\frac{3NP\hbar}{(2I+1)k\Theta}x(2m_1-1)\\ &\times\{\sin(x\xi)\cos^2(x\xi/2)\sin\omega t\cdot\exp(-t^2/2T_2^{*2})\\ &-\sin(x\xi)\sin^2(x\xi/2)\sin\omega(t-2\tau)\\ &\cdot\exp[-(t-2\tau)^2/2T_2^{*2}]+\sin(x\xi)\cos(x\xi)\\ &\times\sin\omega(t-\tau)\cdot\exp[-(t-\tau)^2/2T_2^{*2}]\},\end{aligned}$$

(two pulses). (18)

In Eq. (28), the first and third terms are easily identified with the free induction signals following the first and second pulses respectively while the second term gives an echo at $t=2\tau$. We have here assumed that the rf field is perpendicular to the symmetry axis in the crystal. But in the general case, the rf coil may make an angle θ_1 with the symmetry axis. In such a case, following the above procedure, first it may be shown that in the R-matrix we have now $\xi=\omega_1 t_w\sin\theta_1$, and secondly

the component of the magnetization contributing to the free induction or echo signals will now be proportional to $\langle I_x\rangle \sin\theta_1$. With these modifications the spin-echo amplitude will still be given by (28). These considerations bring our results in exact agreement with those of Bloom[3] and Norberg and Hahn[2] and Herzog, if we put $I=3/2$ and $m_1=3/2$. Thus for spin 3/2, the amplitude of the spin-echo signal is proportional to

$$(\sqrt{3}/2)\sin(\sqrt{3}\xi)\sin^2(\sqrt{3}\xi/2).$$

For a higher spin like 2 or 5/2, where we have two possible quadrupole resonance frequencies, the spin-echo amplitude to be expected for either frequency is easily found by putting down the corresponding values of m_1 and x in Eq. (28). In Table I, we have tabulated the values of the different resonance frequencies, and relative spin-echo amplitudes for spins 1, 3/2, 2, 5/2 due to pure quadrupole coupling alone.

SECTION II. STRONG QUADRUPOLE COUPLING AND WEAK MAGNETIC FIELD

Suppose now we apply a weak steady magnetic field H_0 to the crystal at an angle θ_0 to the axis of symmetry; the Hamiltonian in the absence of the rf field will now be

$$\mathcal{H}=\mathcal{H}_0+\mathcal{H}', \tag{29}$$

where

$$\begin{aligned} &H_0=P\hbar(3I_0^2-\mathbf{I}^2),\\ &\mathcal{H}'=-\omega_0\hbar[I_0\cos\theta_0+\tfrac{1}{2}(I_++I_-)\sin\theta_0], \qquad (30)\\ &\omega_0=\gamma H_0. \end{aligned}$$

It is now easy to see, using the perturbation theory for degenerate states, that the coincidence of the levels $+|m|$ and $-|m|$ discussed in Sec. I is destroyed and the energy levels are characterized by m, each m level being now quite distinct. For half-integral spins however, the two lowest levels are obtained by a mixing of the states corresponding to $m=\pm\frac{1}{2}$. The energy levels and eigenstates are tabulated in Figs. 1 (B and D) respectively for half-integral and integral spins.

We have

$$\begin{aligned} &|+\rangle=|\tfrac{1}{2}\rangle\cos\theta+|-\tfrac{1}{2}\rangle\sin\theta,\\ &|-\rangle=+|-\tfrac{1}{2}\rangle\cos\theta-|\tfrac{1}{2}\rangle\sin\theta, \qquad (31)\\ &\sin\theta=-[(f+1)/2f]^{\frac{1}{2}},\quad \cos\theta=[(f-1)/2f]^{\frac{1}{2}}. \end{aligned}$$

For a half-integral spin $I=\frac{1}{2}(2p+1)$,

$$f=[1+(p+1)^2\tan^2\theta_0]^{\frac{1}{2}}. \tag{32}$$

We shall designate the new eigenstates by the numbers n, with the understanding that the n's correspond to the m's for all the levels of integral spins and for all the levels of half-integral spins, except $m=\pm\frac{1}{2}$. For this lowest pair of levels for half-integral spins, the n's correspond to $+$ and $-$.

Our discussion in this section will be divided into two parts. In the first part we shall take up the case of transitions between the lowest levels of half-integral spins and the next higher. In the second part we shall consider transition between levels other than the lowest for half-integral spins, and between all successive levels of integral spins, i.e., between levels where there is correspondence between the n's and m's.

(a) Transition between the Levels ± and the Next Higher Levels

The states next higher than the lowest levels will be those corresponding to $m=\pm 3/2$. It will be directly seen that the single frequency $\omega(3/2)=6P$ of the pure quadrupole case, is now replaced by four closely spaced frequencies given by

$$\omega_{1,2,3,4}=6P\pm\tfrac{1}{2}(3\pm f)\omega_0\cos\theta_0. \tag{33}$$

Let the applied radio-frequency be $\omega=6P$; to a first approximation we can then assume resonance conditions to exist for all these four frequencies. The R-matrix may be calculated as in Sec. I, with the following modifications. We have to replace $\mathcal{H}_0$ in Eq. (14) by $\mathcal{H}$ given by (29) and also we have to take matrix elements with the eigenstates of $\mathcal{H}$ as basis and not those of $\mathcal{H}_0$. A similar modification has to be applied to Eq. (18) to obtain the D-matrix in the present case. Also, exactly as in Sec. I, we can obtain in the present case the relevant matrix elements of S_{I} and S_{II}, the echo matrices after one and two pulses respectively. We can then use (20)[15] to find the expectation value of I_x after one and two pulses respectively to obtain the free-induction and echo amplitudes. We have the following equations for the expectation value of I_x after

TABLE I. Spin-echo amplitudes in pure quadrupole coupling case for $I=1, 3/2, 2, 5/2$.

Spin	Value of $P(I)$	Resonance frequencies	Spin-echo amplitudes
1	$\frac{e^2Qq}{\hbar}$	$3P(1)$	$\frac{NP(1)\hbar}{k\Theta}\sqrt{2}\sin(\sqrt{2}\xi)\sin^2(\xi/\sqrt{2})$
3/2	$\frac{e^2Qq}{12\hbar}$	$6P(3/2)$	$\frac{3NP(3/2)\hbar}{4k\Theta}2\sqrt{3}\sin(\sqrt{3}\xi)\sin^2\left(\frac{\sqrt{3}\xi}{2}\right)$
		$9P(2)$	$\frac{3NP(2)\hbar}{5k\Theta}6\sin(2\xi)\sin^2\xi$
2	$\frac{e^2Qq}{24\hbar}$		
		$3P(2)$	$\frac{3NP(2)\hbar}{5k\Theta}\sqrt{6}\sin(\sqrt{6}\xi)\sin^2\left(\frac{\sqrt{6}\xi}{2}\right)$
		$12P(5/2)$	$\frac{3NP(5/2)\hbar}{6k\Theta}4\sqrt{5}\sin(\sqrt{5}\xi)\sin^2\left(\frac{\sqrt{5}\xi}{2}\right)$
5/2	$\frac{e^2Qq}{40\hbar}$		
		$6P(5/2)$	$\frac{3NP(5/2)\hbar}{6k\Theta}2\sqrt{8}\sin(\sqrt{8}\xi)\sin^2(\sqrt{2}\xi)$

[15] Properly we should use (4) with $\rho(0)$ defined by $\rho(0)=\exp(-\mathcal{H}/k\Theta)$, with $\mathcal{H}$ given by (29), but if we do not want to retain terms of the order of ω_0/P in the free-induction signal and echo amplitudes, we may just as well use (20) without much error.

one and two pulses, *viz.*,

$$\langle I_x\rangle=\frac{3NP\hbar}{k\Theta}\frac{2x}{2I+1}\sin(x\xi)\sin[\omega(\tfrac{3}{2})\cdot t]\cdot\left[\frac{f+1}{f}\cos\left(\frac{3-f}{2}(\omega_0\cos\theta_0)t\right)+\frac{f-1}{f}\cos\left(\frac{3+f}{2}(\omega_0\cos\theta_0)t\right)\right]$$

(single pulse), (34)

$$\langle I_x\rangle=\frac{3NP\hbar}{k\Theta}\frac{2x}{2I+1}\sin(x\xi)\sin^2\left(\frac{x\xi}{2}\right)\sin[\omega(\tfrac{3}{2})\cdot(t-2\tau)]\times\left\{\left(\frac{f+1}{2f}\right)^2\cos\left(\frac{3-f}{2}(\omega_0\cos\theta_0)(t-2\tau)\right)+\left(\frac{f-1}{2f}\right)^2\cos\left(\frac{3+f}{2}(\omega_0\cos\theta_0)(t-2\tau)\right)+\frac{f^2-1}{2f^2}\left[\cos\left(\frac{3\omega_0\cos\theta_0}{2}(t-2\tau)\right)\cos\left(\frac{f\omega_0\cos\theta_0}{2}t\right)+\cos\left(\frac{3\omega_0\cos\theta_0}{2}t\right)\cos\left(\frac{f\omega_0\cos\theta_0}{2}(t-2\tau)\right)-\cos\left(\frac{3\omega_0\cos\theta_0}{2}t\right)\cos\left(\frac{f\omega_0\cos\theta_0}{2}t\right)\right]\right\}$$

(two pulses). (35)

In Eq. (35), only the terms giving rise to an echo at $t=2\tau$, *viz.*, the terms involving $\sin[\omega(3/2)\cdot(\kappa-\tau)]$ are given. In the presence of an inhomogeneity in q and a corresponding inhomogeneity in P, we have to integrate over a Gaussian distribution in $\Delta\omega=\omega(3/2)-\omega$ with rms value equal to $1/T_2^*$ as before. $\sin(\omega(3/2)\cdot t)$ in (34) then gets replaced by $\sin\omega t\cdot\exp(-t^2/2T_2^{*2})$ and $\sin[\omega(3/2)\cdot(t-2\tau)]$ in (35) gets replaced by $\sin(\omega(t-2\tau))\cdot\exp[-(t-2\tau)^2/2T_2^{*2}]$, the latter exponential term showing the occurrence of the echo maximum at $t=2\tau$. It is evident that we get a modulation of the pure-quadrupole echo signal by the action of the weak z-field. If only a single spin-echo signal, corresponding to a pair of pulses with a fixed interval between them, be observed, then it will appear to have a modulation on account of all the terms in (35) involving ω_0. We term this modulation pattern "echo modulation." On the other hand, in echo envelope measurements with paired pulses of varying intervals τ only the last ω_0 terms, *viz.*, those multiplying the factor $(f^2-1)/2f^2$ will give rise to a modulation of the envelope, the other two terms being independent of τ at the instants of occurrence of the echoes, i.e., at $t=2\tau$. We term this modulation "envelope modulation." Besides this modulation of the echo envelope there will be the usual exponential damping due to the causes contributing to the transverse relaxation time. We shall not incorporate the effect of the latter here, but shall merely mention that as pointed out by Hahn and Maxwell,[5] to get a correct estimate of the present modulation effects from experimental echo envelopes, the latter have to be normalized first to correct for the above damping.

(b) The other important case that we have to consider is that in which the applied frequency corresponds to pure quadrupole frequencies $\omega(m_1)$ other than that between levels $\pm3/2$ and $\pm1/2$. In this case, each frequency $\omega(m_1)$ in the pure quadrupole case is broken up into a doublet $\omega(m_1)\pm\omega_0\cos\theta_0$ by the action of the weak magnetic field. Proceeding as in the previous two cases, we obtain the free induction and echo signals respectively as

$$\langle I_x\rangle=\frac{3NP\hbar}{k\Theta}\frac{x(2m_1-1)}{(2I+1)}\sin(x\xi)\sin\omega t\times\cos[\omega_0\cos\theta_0\cdot t]\exp\left(-\frac{t^2}{2T_2^{*2}}\right),$$

$$\langle I_x\rangle=\frac{3NP\hbar}{k\Theta}\frac{x(2m_1-1)}{(2I+1)}\sin(x\xi)\sin^2\left(\frac{x\xi}{2}\right)\sin[\omega(t-2\tau)]\times\cos[\omega_0\cos\theta_0(t-2\tau)]\times\exp[-(t-2\tau)^2/2T_2^{*2}]. \quad (36)$$

This expression evidently holds for all the pure quadrupole resonance frequencies of a nucleus with integral spin I and for all those of half-integral spins except the lowest one, *viz.*, that between levels $\pm3/2$ and $\pm1/2$. It is evident from the expression (36), that in these cases we merely have an "echo modulation" and not any "envelope modulation."

SECTION III. STRONG STEADY MAGNETIC FIELD AND WEAK QUADRUPOLE COUPLING

In this case, the Hamiltonian in the absence of the rf field is given by

$$\mathcal{H}=\mathcal{H}_0+\mathcal{H}', \quad (37)$$

where

$$\mathcal{H}_0=-\gamma\hbar I_z H_z,$$

and

$$\mathcal{H}'=\frac{e^2qQ}{8I(2I-1)}[(3I_0^2-\mathbf{I}^2)(3\cos^2\theta_0-1)+\tfrac{3}{2}\sin\theta_0\cos\theta_0(e^{-i\phi_0}\{I_+,I_0\}+e^{i\phi_0}\{I_-,I_0\})+\tfrac{3}{4}\sin^2\theta_0(I_+^2e^{-2i\phi_0}+I_-^2e^{2i\phi_0})], \quad (38)$$

Q and q being defined as before, and θ_0 giving the angle between the symmetry axis and the z-direction, *viz.*, the direction of the steady field; $\{a,b\}$ represents the anticommutator $(ab+ba)$ as before. We then have the

energy levels correct to the first order given by

$$E_m=-m\hbar\omega_z+\tfrac{1}{2}P\hbar(3\cos^2\theta_0-1)[3m^2-I(I+1)], \quad (39)$$

the eigenstates being characterized by the magnetic quantum number m, and

$$P=\frac{e^2qQ}{4I(2I-1)\hbar}. \quad (9)$$

This case differs from the previous cases in that there is only one central frequency here, *viz.*, ω_z, about which we have a number of symmetrically placed components with frequency separations of the order of P. Hence all these frequencies will be more or less subjected to resonance simultaneously and so the R-matrix will no longer be of the simple type obtained in the preceding cases. There will now be matrix elements between all the levels and it thus becomes difficult to give a generalized treatment for spin I. We shall only give the results for a few important cases, *viz.*, $I=1$, $3/2$, 2. The last one is important, because whereas in the first two cases there is only one symmetrical pair of lines, in this case there are two. The case of spins greater than $I=2$ may be worked out by following the general procedure outlined below.

The R-matrix may be found as usual by Eq. (14), employing the eigenvalues and eigenkets discussed above. We then find that if we agree to neglect terms of the order of P/ω in the amplitude of the spin-echo term, then the R-matrix is the same as the usual Rabi-Bloch matrix,[16] *viz.*,

$$\langle m|R|m'\rangle=\exp[it_w\{m\omega_z-P[3m^2-I(I+1)]\times(3\cos^2\theta_0-1)/2\}]\cdot\langle m|R^*|m'\rangle,$$

with

$$\langle m|R^*|m'\rangle=\{(I+m)!(I-m)!(I+m')!(I-m')!\}^{\frac{1}{2}}\times\sin^{2I}(\xi/2)i^{(2I-m-m')}\times\sum_k\frac{(-)_k\cot^{(m+m'+2k)}(\xi/2)}{(m+m'+k)!(I-m-k)!(I-m'-k)!k!}, \quad (40)$$

where k takes up integral values restricted by the condition that none of the factorial terms in the denominator is to be negative.

The D-matrix will as usual be given by

$$D(t,t_0)=\exp[-(i/\hbar)\mathcal{H}(t-t_0)]. \quad (41)$$

The density matrix at start will be given by the Boltzmann distribution, *viz.*,

$$\rho(0)=\frac{N}{2I+1}\exp\left(-\frac{\mathcal{H}}{k\Theta}\right),$$

and if we are not interested in retaining terms of the order P/ω in the echo and free-induction signal amplitudes, we can safely take it as

$$\rho(0)=\frac{N}{2I+1}\exp\left(-\frac{\hbar\omega_zI_z}{k\Theta}\right). \quad (42)$$

We can now obtain the echo matrices S_{I} and S_{II} after one and two pulses respectively, using (40), (41) and (2). Using these matrices and Eq. (3), we get the density matrix after the passage of one and two pulses respectively. Using (4), we now obtain the following expressions for the free induction and echo signals[17] for spin $I=1$:

Free-induction signal:

$$\langle I_x\rangle=\frac{2N\hbar\omega_z}{(2I+1)k\Theta}\sin\xi\sin\omega_zt\cos(3Pt),$$

Spin-echo signal:

$$\langle I_x\rangle=\frac{2N\hbar\omega_z}{(2I+1)k\Theta}\sin\xi\sin^2(\xi/2)\sin[\omega_z(t-2\tau)]\times\{\cos\xi\cos(3Pt)-2\cos^2(\xi/2)\cos[3P(t-2\tau)]\}, \quad (43)$$

where $\omega_z\pm3P$ are the two frequencies into which the original resonance line due to the z-field alone is split up by the quadrupole interaction. We have to integrate these expressions over the inhomogeneity in the applied z-field as in Paper I; we then get the free-induction and spin-echo signals as

Free-induction signal:

$$\langle I_x\rangle=-\frac{2}{3}\frac{N\hbar\omega}{k\Theta}\sin\xi\cos(3Pt)\cdot\exp\left(-\frac{t^2}{2T_2^{*2}}\right), \quad (44)$$

Spin-echo signal:

$$\langle I_x\rangle=-\frac{2}{3}\frac{N\hbar\omega}{k\Theta}\sin\xi\sin^2(\xi/2)\cdot\{\cos\xi\cos3Pt-2\cos^2(\xi/2)\cos[3P(t-2\tau)]\}. \quad (45)$$

It is evident that in the echo-expression (45), only the first term in the brackets causes an "envelope-modulation." The resonance frequencies and the free induction and echo signal amplitudes for spins $I=1$, $3/2$ and 2 have been tabulated in Table II. A check on the correctness of these echo amplitudes is obtained by putting $P=0$, when we get the expression $\frac{1}{3}I(I+1)\times(2I+1)\sin\xi\sin^2(\xi/2)$ for spin I, discussed in Paper I.

CONCLUSION

We have discussed above the two extreme cases of spin-echo formation in crystals with (a) strong quadru-

[16] F. Bloch and I. I. Rabi, Revs. Modern Phys. 17, 237 (1945).

[17] We have only given the results for the case when the symmetry axis of the axially symmetric electric field gradient tensor, at the sites of the resonant nuclei, coincides with the direction of the applied magnetic field, i.e., $\theta_0=0$.

TABLE II. Spin-echo amplitudes in the case of strong z-field and weak quadrupole coupling.

Spin	Value of $P(I)$	Resonance frequencies	Free-induction signal amplitude[a]	Echo signal amplitude[a]
1	$\frac{e^2Qq}{4h}$	$\omega_z \pm 3P(1)$	$\frac{2}{3}\frac{N\hbar\omega}{k\Theta}S\cos(3P(1)t)$	$\frac{4}{3}\frac{N\hbar\omega}{k\Theta}S^3C\{(C^3-S^3)\cos(3Pt)-2C^3\cos[3P(t-2\tau)]\}$
3/2	$\frac{e^2Qq}{12h}$	$\omega_z \pm 6P(3/2)$	$\frac{N\hbar\omega}{k\Theta}SC[1+\frac{3}{2}\cos(6P(3/2)t)]$	$\frac{1}{2}\frac{N\hbar\omega}{k\Theta}S^3C\{-2(2C^2-S^2)^2+6C^2(C^2-2S^2)\cos(6P\tau)$ $-9C^4\cos[6P(t-2\tau)]+3S^2(2C^2-S^2)\cos(6Pt)$ $+6C^2(C^2-2S^2)\cos[6P(t-\tau)]\}$
2	$\frac{e^2Qq}{24h}$	$\omega_z \pm 9P(2)$ $\omega_z \pm 3P(2)$	$\frac{4N\hbar\omega}{5k\Theta}SC[2\cos 9Pt+3\cos 3Pt]$	$\frac{2}{5}\frac{N\hbar\omega}{k\Theta}S^3C\{-16C^6\cos 9P(t-2\tau)$ $-4S^4(S^2-29C^4+86C^6S^2+20C^8)\cos(9Pt)$ $+12C^6(C^2-3S^2)(C^4+S^4)\cos[3P(3t-4\tau)]$ $+24C^4S^2[2S^2(S^2-C^2)+C^4(2+C^2)]\cos[3P(3t-2\tau)]$ $+12C^4[C^8-C^4S^2-11S^4C^2+17S^6-24S^8]\cos 3P(t-4\tau)$ $-36C^2(C^8-2S^4C^4+6S^6C^2-5S^8+4S^{10})\cos 3P(t-2\tau)$ $+12(-S^8+7S^4C^2-13S^2C^4+3C^6)\cos(3Pt)$ $+24S^2C^2(C^6-C^4S^2+C^2S^4-3S^6+2S^8)\cos[3P(t+2\tau)]\}$

[a] C and S denote $\cos(\xi/2)$ and $\sin(\xi/2)$ respectively. In some places, due to lack of space the brackets denoting the spin to which P refers have been omitted. In these cases, the P's naturally stand for the $P(I)$'s of the corresponding rows.

pole interaction and weak magnetic field and (b) weak quadrupole interaction and strong magnetic field. Experiments in the former case have already been reported. But there are as yet no experimental results in the latter case. In some cases, such as in Al^{27} resonance in Al_2O_3 crystal, as Pound[9] points out, the second order perturbation due to electric quadrupole interaction is fairly appreciable. In such cases we expect some departure from the results obtained by the method detailed above. Nevertheless, using a second order perturbation treatment to handle the electric quadrupole interaction, we can obtain the spin-echo amplitudes in these cases, using the rudiments of our method. Besides, we have only discussed the situation when the strong magnetic field and the symmetry axis of the crystal coincide. We shall subsequently publish the results of calculation for the case in which the symmetry axis of the crystal makes an angle θ_0 with the strong magnetic field. Of course, the most general case would be one of "intermediate coupling," where the quadrupole interaction and the interaction with the z-field are of the same order. In this case, as Feld and Lamb[18] have discussed, we have to solve a secular equation to obtain the energy levels and eigenkets of the total Hamiltonian (29), the two terms $\mathcal{H}_0$ and $\mathcal{H}'$ being now of the same order. A complete solution of this secular equation (usually cubic, for spin 1 and higher than cubic for spins $I>1$) can only be obtained numerically. Using the energy values and eigenstates so obtained we can then find the R and D matrices using Eqs. (14) and (18). Then, proceeding as in the previous cases, the free-induction and echo amplitudes may be found by evaluating $\langle I_z \rangle$ after one and two pulses respectively.

The authors are indebted to Professor M. N. Saha for his keen interest in the progress of the work. They are also grateful to Mr. D. K. Roy and Mr. M. K. Banerjee for some helpful discussions.

[18] B. T. Feld and W. E. Lamb, Phys. Rev. **67**, 15 (1945).

Spin Echoes with Four Pulses—An Extension to n Pulses

T. P. Das and D. K. Roy
Institute of Nuclear Physics, Calcutta, India

(Received June 7, 1954; revised manuscript received August 12, 1954)

Following Das and Saha's method, the response of a nuclear spin system to four rf pulses has been analyzed by using the classical Bloch equations. The calculations enable one to obtain a recursion formula for the number of echoes to be expected with n pulses. Besides, the calculations also yield several other results important from the experimental view. Thus, they provide a method for obtaining T_1, the spin-lattice relaxation time, from primary echo measurements alone, and also give an idea of the order of correction necessary in a measurement of T_2, the "spin-spin relaxation time," by observation of primary echoes with successive paired pulses of gradually changing intervals.

I. INTRODUCTION

HAHN[1] first gave the mathematical theory for the occurrence of free-induction and spin-echo signals when the rf field is applied in the form of short pulses. Das and Saha[2] subsequently showed that in estimating the damping effects of diffusion on these signals, the frequency difference and phase shift accumulated by a certain "isochromatic" group in any free-precession interval, by virtue of diffusion, could be regarded as independent of those accumulated in the previous intervals. Hence they can be averaged independently after collecting all the terms that are of even parity in all the variables involved. We have followed this procedure in our calculations with four pulses, the results of which we present in Sec. II. In Sec. III, we have interpreted the various terms obtained in Sec. II in terms of the primary and stimulated echo mechanisms between applied pulses and between echoes and applied pulses. One term remains unexplained after applying this mode of explanation. It appears to be characteristic of four pulses alone and follows a new law of combination of intervals as regards its position. We call it a quaternary echo; subsequent calculations with five and six pulses also show that we have similar "pentanary" and "hexanary" echoes occurring with five and six pulses respectively. Using this idea, we establish a recursion formula for the number of echoes to be expected in general with n pulses. In Sec. IV, we discuss the corrections to be applied to primary and stimulated echo envelope measurements with successive groups of two and three pulses respectively when the intervals between these successive groups are only fractions of T_1. A possible method for obtaining T_1, by measurements on primary echo envelopes alone, is suggested.

II. RESULTS WITH FOUR PULSES

As shown in Fig. 1, we take four rf pulses of equal width. The angle of nutation during each pulse is therefore given by

$$\xi=\omega_1 t_w.$$

We follow the same terminology and notation as in I. Thus, we denote the frequency drifts undergone by the constituents of a certain isochromatic group in the first three free-precession intervals by η_{10}, η_{21}, and η_{32} respectively, while we denote the phase shifts in these three intervals and in the interval following the fourth pulse by ϕ_{10}, ϕ_{21}, ϕ_{32}, and ϕ_{43} respectively. The distribution functions for these are given by Eqs. (31) and (33) of I. Also, for completeness, and to enable us to draw conclusions about the number of echoes to be expected with n pulses, we assume that the following conditions hold regarding the lengths of the different free-precession intervals:

$$\tau_3>2\tau_2 \quad \text{and} \quad \tau_2>2\tau_1. \tag{1}$$

If these conditions do not hold, then as we shall presently obtain by our calculations, some of the echoes will not be observed.

To obtain the conditions after the passage of the fourth pulse, we have to proceed as in I, by solving the Bloch equations during and after a pulse, and tabulating the solutions at the different stages of passage of pulses. For our present purpose, we shall need, besides the solutions tabulated in I, *viz.*, those at A, B, C, D, F, G, and H, also those at I, J, and K. Following an exactly similar procedure as in I, we can then obtain the value of V after the passage of four pulses, *viz.*, $V(K)$. It will be seen to consist of 22 terms, whose positions and the different parts of the amplitudes, *viz.*, trigonometric, diffusion damping, and relaxation damping parts are tabulated in Table I. The terminology for the different terms will be explained in Sec. III.

III. DISCUSSION OF THE VARIOUS TERMS IN TABLE I; THE RECURSION FORMULA FOR n PULSES

We shall now explain the terminology used for the different terms in Table I and interpret them physically.

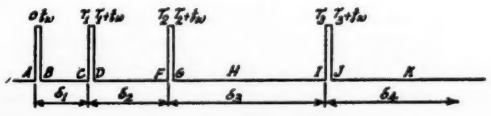


FIG. 1. Applied rf pulses.

[1] E. L. Hahn, Phys. Rev. **80**, 580 (1950).
[2] T. P. Das and A. K. Saha, Phys. Rev. **93**, 749 (1954). We shall refer to this paper henceforth as I.

These terms can evidently be divided into the following categories.

(*a*) *Free induction terms.*—The symbol used for any one of these is in general F_n, representing the free induction signal following the nth pulse. In the case of four applied pulses there are thus four such terms designated by F_1, F_2, F_3, and F_4.

(*b*) *Terms arising from a primary echo mechanism between any two of the applied pulses.*—A typical one is denoted by $P_{(mn)}$, which represents the primary echo arising from a pair of rf pulses respectively applied at instants τ_m and τ_n. The interval between the first pulse of the pair and the echo is $2(\tau_n-\tau_m)=2\delta_{nm}$, δ_{nm} representing the interval between the pair of pulses. There

TABLE I. Amplitudes of the different free induction and echo terms for four pulses.[a]

Term	Position of maximum	Trigonometric part of amplitude	Diffusion damping of the amplitude[b]	Relaxation damping of the amplitude
F_1	0	$-\sin\xi\cos^6(\xi/2)$	$\exp\left[-\frac{k}{3}\left(f_1+\sum_r^{1,4}\delta_r^3\right)\right]$	$\exp\left(-\frac{t}{T_2}\right)$
F_2	τ_1	$-W(C)\sin\xi\cos^4(\xi/2)$	$\exp\left[-\frac{k}{3}\left(f_1+\sum_r^{2,4}\delta_r^3\right)\right]$	$\exp\left(-\frac{t-\tau_1}{T_2}\right)$
F_3	τ_2	$-\sin\xi\sin^2(\xi/2)$	$\exp\left[-\frac{k}{3}\left(f_2+\sum_r^{3,4}\delta_r^3\right)\right]$	$\left\{1-[1-W(C)\cos\xi]\exp\left(-\frac{\delta_2}{T_1}\right)\right\}\times\exp\left(-\frac{t-\tau_2}{T_2}\right)$
F_4	τ_3	$\sin\xi$	$\exp\left[-\frac{k}{3}(f_3+\delta_4^3)\right]$	$\left\{1-(1-\cos\xi)\exp\left(-\frac{\delta_3}{T_1}\right)-\cos\xi[1-W(C)\cos\xi]\times\exp\left(-\frac{\tau_3-\tau_1}{T_1}\right)\right\}\exp\left(-\frac{t-\tau_3}{T_2}\right)$
$P_{(12)}$	$2\tau_1$	$\frac{1}{4}\sin^3\xi\cos^2(\xi/2)$	$\exp\left[-\frac{k}{3}\left(f_1+\sum_r^{1,4}\delta_r^3\right)\right]$	$\exp\left(-\frac{t}{T_2}\right)$
$P_{(13)}$	$2\tau_2$	$\frac{1}{4}\sin^3\xi\cos^2(\xi/2)$	$\exp\left[-\frac{k}{3}\left(f_4+\sum_r^{1,4}\delta_r^3\right)\right]$	$\exp\left(-\frac{t}{T_2}\right)$
$P_{(23)}$	$2\tau_2-\tau_1$	$\frac{1}{4}W(C)\sin^3\xi$	$\exp\left[-\frac{k}{3}\left(f_4+\sum_r^{2,4}\delta_r^3\right)\right]$	$\exp\left(-\frac{t-\tau_1}{T_2}\right)$
$P_{(14)}$	$2\tau_3$	$\frac{1}{4}\sin^3\xi\cos^2(\xi/2)$	$\exp\left[-\frac{k}{3}\left(f_5+\sum_r^{1,4}\delta_r^3\right)\right]$	$\exp\left(-\frac{t}{T_2}\right)$
$P_{(24)}$	$2\tau_3-\tau_1$	$\frac{1}{4}W(C)\sin^3\xi$	$\exp\left[-\frac{k}{3}\left(f_5+\sum_r^{2,4}\delta_r^3\right)\right]$	$\exp\left(-\frac{t-\tau_2}{T_2}\right)$
$P_{(34)}$	$2\tau_3-\tau_2$	$\sin\xi\sin^2(\xi/2)$	$\exp\left[-\frac{k}{3}\left(f_6+\sum_r^{3,4}\delta_r^3\right)\right]$	$\left\{1-[1-W(C)\cos\xi]\exp\left(-\frac{\delta_2}{T_1}\right)\right\}\times\exp\left(-\frac{t-\tau_2}{T_1}\right)$

[a] Here, $\delta_1=\tau$, $\delta_2=\tau_2-\tau_1$, $\delta_3=\tau_3-\tau_2$, $\delta_4=t-\tau_3$,

$f_1=3(t-\tau_1)^2\tau_1+3(t-\tau_2)^2(\tau_2-\tau_1)+3(t-\tau_3)^2(\tau_3-\tau_2)$,
$f_2=3(t-\tau_2)^2\tau_1+3(t-\tau_2)^2(\tau_2-\tau_1)+3(t-\tau_3)^2(\tau_3-\tau_2)$,
$f_3=3(t-\tau_3)^2\tau_1+3(t-\tau_3)^2(\tau_3-\tau_1)+3(t-\tau_3)^2(\tau_3-\tau_2)$,
$f_4=3(t-2\tau_2+\tau_1)^2\tau_1+3(t-\tau_2)^2(\tau_2-\tau_1)+3(t-\tau_3)^2(\tau_3-\tau_2)$,
$f_5=3(t-2\tau_3+\tau_1)^2\tau_1+3(t-2\tau_3+\tau_2)^2(\tau_2-\tau_1)+3(t-\tau_3)^2(\tau_3-\tau_2)$,
$f_6=3(t-2\tau_3+\tau_2)^2\tau_1+3(t-2\tau_3+\tau_2)^2(\tau_2-\tau_1)+3(t-\tau_3)^2(\tau_3-\tau_2)$,
$f_7=3(t-2\tau_3+2\tau_2-\tau_1)^2\tau_1+3(t-2\tau_3+\tau_2)^2(\tau_2-\tau_1)+3(t-\tau_3)^2(\tau_3-\tau_2)$,
$f_8=3(t-\tau_3-\tau_2+\tau_1)^2\tau_1+3(t-\tau_2)^2(\tau_2-\tau_1)+3(t-\tau_3)^2(\tau_3-\tau_2)$.

[b] By $\sum_r^{m,n}\delta_r^3$ we mean $\delta_m^3+\delta_{m+1}^3+\cdots+\delta_n^3$.

TABLE I (*continued*).

Term	Position of maximum	Trigonometric part of amplitude	Diffusion damping of the amplitude[b]	Relaxation damping of the amplitude
$P_{((12)3)}$	$2\tau_2-2\tau_1$	$-\frac{1}{4}\sin^3\xi\sin^2(\xi/2)$	$\exp\left[-\frac{k}{3}\left(f_4+\sum_r^{1,4}\delta_r^3\right)\right]$	$\exp\left(-\frac{t}{T_2}\right)$
$P_{((12)4)}$	$2\tau_3-2\tau_1$	$-\frac{1}{4}\sin^3\xi\sin^2(\xi/2)$	$\exp\left[-\frac{k}{3}\left(f_5+\sum_r^{1,4}\delta_r^3\right)\right]$	$\exp\left(-\frac{t}{T_2}\right)$
$P_{((13)4)}$	$2\tau_3-2\tau_2$	$-\frac{1}{4}\sin^3\xi\sin^2(\xi/2)$	$\exp\left[-\frac{k}{3}\left(f_7+\sum_r^{1,4}\delta_r^3\right)\right]$	$\exp\left(-\frac{t}{T_2}\right)$
$P_{((22)4)}$	$2\tau_3-2\tau_2-\tau_1$	$-W(C)\sin\xi\sin^4(\xi/2)$	$\exp\left[-\frac{k}{3}\left(f_7+\sum_r^{2,4}\delta_r^3\right)\right]$	$\exp\left(-\frac{t-\tau_1}{T_2}\right)$
$P_{(((12)3)4)}$	$2\tau_3-2\tau_2+2\tau_1$	$\sin\xi\sin^6(\xi/2)$	$\exp\left[-\frac{k}{3}\left(f_7+\sum_r^{1,4}\delta_r^3\right)\right]$	$\exp\left(-\frac{t}{T_2}\right)$
$P_{(123)}$	$\tau_2+\tau_1$	$\frac{1}{2}\sin^3\xi\sin^2(\xi/2)$	$\exp\left[-\frac{k}{3}(f_3+\delta_1^3+\delta_2^3+\delta_4^3)\right]$	$\exp\left(-\frac{t-\tau_2+\tau_1}{T_2}-\frac{\delta_2}{T_1}\right)$
$P_{(134)}$	$\tau_3+\tau_2$	$-\frac{1}{2}\sin^3\xi\cos^2(\xi/2)$	$\exp\left[-\frac{k}{3}(f_3+\delta_1^3+\delta_2^3+\delta_4^3)\right]$	$\exp\left(-\frac{t-\tau_3+\tau_2}{T_2}-\frac{\delta_3}{T_1}\right)$
$P_{(124)}$	$\tau_3+\tau_1$	$-\frac{1}{2}\sin^3\xi\cos\xi$	$\exp\left[-\frac{k}{3}(f_3+\delta_1^3+\delta_4^3)\right]$	$\exp\left(-\frac{t-\tau_3+\tau_1}{T_2}-\frac{\tau_3-\tau_1}{T_1}\right)$
$P_{(234)}$	$\tau_3+\tau_2-\tau_1$	$-\frac{1}{4}W(C)\sin^3\xi$	$\exp\left[-\frac{k}{3}(f_3+\delta_2^3+\delta_4^3)\right]$	$\exp\left(-\frac{t-\tau_3+\tau_2-\tau_1}{T_2}-\frac{\delta_2}{T_1}\right)$
$P_{((123)4)}$	$2\tau_3-(\tau_2+\tau_1)$	$-\frac{1}{2}\sin^3\xi\cos^2(\xi/2)$	$\exp\left[-\frac{k}{3}(f_7+\delta_1^3+\delta_2^3+\delta_4^3)\right]$	$\exp\left(-\frac{t-\tau_2+\tau_1}{T_2}-\frac{\delta_2}{T_1}\right)$
$P_{((12)34)}$	$\tau_3+\tau_2-2\tau_1$	$\frac{1}{2}\sin^3\xi\sin^2(\xi/2)$	$\exp\left[-\frac{k}{3}(f_3+\delta_1^3+\delta_2^3+\delta_4^3)\right]$	$\exp\left(-\frac{t-\tau_3+\tau_2}{T_2}-\frac{\delta_3}{T_1}\right)$
$P_{(1234)}$	$2\tau_3-\tau_2+\tau_1$	$-\frac{1}{2}\sin^3\xi\cos^2(\xi/2)$	$\exp\left[-\frac{k}{3}(f_3+\delta_1^3+\delta_3^3+\delta_4^3)\right]$	$\exp\left(-\frac{t-\tau_2+\tau_1}{T_2}-\frac{\delta_2}{T_1}\right)$

are six such echoes in the case of four applied pulses as seen from Table I, and the above law for finding their positions is seen to give their positions correctly.

(*c*) *Terms which can be explained by a primary-echo mechanism between primary echoes due to previous applied rf pulses and subsequent rf pulses.*—A typical one is denoted by $P_{((mn)p)}$, which gives the primary echo due to primary-echo interaction between the primary echo $P_{(mn)}$ and the applied pulse at τ_p. The interval between the echo $P_{((mn)p)}$ and the echo $P_{(mn)}$ is given by $2\delta_{p,\,(mn)}$, where $\delta_{p,\,(mn)}$ is the interval between the applied rf pulse at τ_p and the primary echo $P_{(mn)}$, the position of the latter being found by the combination rule given in the previous paragraph. This rule gives correctly the positions of the maxima of four such terms occurring in Table I with four applied rf pulses. There is one more echo in Table I which comes into this category, *viz.* $P_{(((12)3)4)}$, which represents the primary echo due to interaction between the primary echo $P_{((12)3)}$ and the fourth applied pulse. We can make a general notation $P_{(((mn)p)q)}$ for this type of echo. The rule for finding its position is that it is at an interval $2\delta_{q,\,((mn)p)}$ away from the echo $P_{((mn)p)}$ with $\delta_{q,\,((mn)p)}$ representing the interval between the pulse at τ_q and the echo $P_{((mn)p)}$; the position of the latter is found by the combination rule given in the first part of this paragraph. A similar interpretation may be given to terms like $P_{((((\cdots(mn)p)\cdots)q)\cdots)}$ which may arise with more than four applied rf pulses.

(*d*) *Terms which may be explained by a "stimulated" echo mechanism between three applied pulses.*—A typical one we denote by $P_{(mnp)}$, arising out of the three applied pulses at τ_m, τ_n and τ_p; the rule for finding its position is that it is situated an interval $(\tau_n-\tau_m)+(\tau_p-\tau_m)$, i.e., $\delta_{pn}+2\delta_{nm}$ away from the first of the trio. There are clearly four such echo terms in Table I, and the above law of combination evidently gives the positions of their maxima correctly.

(*e*) *Terms which can be explained as arising from a primary echo mechanism between a stimulated echo due to three previous rf pulses and a subsequent rf pulse.*—A typical one is denoted by $P_{((mnp)q)}$, arising out of a primary echo—mechanism between the stimulated echo $P_{(mnp)}$ and the applied pulse at τ_q. The law of combination of intervals for finding the position of this echo is that it is an interval $2\delta_{q,(mnp)}$ away from the stimulated echo $P_{(mnp)}$, $\delta_{q,(mnp)}$ representing the interval between τ_q and the echo $P_{(mnp)}$, the position of the latter being given by the rule in the preceding paragraph. There is evidently one such echo in Table I for four applied rf pulses, *viz.* $P_{((123)4)}$, and the above rule is seen to give its position correctly.

(*f*) *Terms which can be explained as arising out of a "stimulated-echo" mechanism between an echo (primary or stimulated) and two subsequent pulses.*—A typical one is denoted by $P_{((lmn)pq)}$ or $P_{((mn)pq)}$, according as the previous echo contributing to the "stimulated echo" mechanism in question is the "stimulated echo" $P_{(lmn)}$ or the primary echo $P_{(mn)}$ respectively, the rule for finding the position being easily found by a combination of the different rules given in the previous paragraphs. We have only one such echo, *viz.* $P_{((12)34)}$ in the case of four pulses tabulated in Table I.

(*g*) *Final term.*—Finally we have one echo which we have denoted by $P_{(1234)}$, with a maximum occurring at $2\tau_3-\tau_2+\tau_1$, which we cannot explain by the mechanism discussed in the above paragraphs. It appears to be due to a new type of combination of intervals which we call "quaternary," characteristic of four pulses, and analogous to the "primary" combination with two pulses, and the "stimulated" combination with three pulses. The rule for finding its position when written in terms of the interval between the pulses is also interesting; it is at an interval $2\delta_3+\delta_2+2\delta_1$ (see Fig. 1) away from the first pulse as compared to $\delta_2+2\delta_1$ for the stimulated and $2\delta_1$ for the primary echo mechanisms. We have also carried out calculations with five and six applied rf pulses, when "pentanary" and "hexanary" mechanisms following the combination rules $\delta_4+2\delta_3+\delta_2+2\delta_1$ and $2\delta_5+\delta_4+2\delta_3+\delta_2+2\delta_1$ are found as well. Hence we find that there is a new law of combination giving rise to a special type of echo, coming in with every extra applied pulse. Of course, all these can be explained by the rotating vector models developed by Hahn.[1,3] In general, the law of combination involving n pulses, giving the interval of the echo in question from the first pulse of the group, will thus be given by

$$\begin{aligned} &2(\delta_{n-1}+\delta_{n-3}+\cdots+\delta_1)+(\delta_{n-2}+\delta_{n-4}+\cdots+\delta_2), \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{for } n \text{ even}, \\ &(\delta_{n-1}+\delta_{n-3}+\cdots+\delta_2)+2(\delta_{n-2}+\delta_{n-4}+\cdots+\delta_1) \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{for } n \text{ odd}. \end{aligned} \tag{2}$$

[3] For a specially simplified picture for explaining the origin of these mechanisms, we can consider rotating models with a 90° pulse followed by 180° pulses, as discussed by H. Y. Carr and E. M. Purcell, Phys. Rev. 94, 630 (1954).

$\delta_1, \delta_2, \cdots$, representing the intervals between successive pulses starting from the first one.

These considerations enable us to obtain a recursion formula for the number of echoes, y_n, to be expected with n pulses.

(*a*) Number of primary echoes due to interaction directly between the nth pulse and any of the previous pulses $=\binom{n-1}{1}$.

(*b*) Number of primary echoes due to interaction between nth pulse and previous echoes $=y_{n-1}+y_{n-2}+\cdots+y_1$.

(*c*) Number of stimulated echoes between nth pulse and any two of the previous pulses $=\binom{n-1}{2}$.

(*d*) Number of stimulated echoes between nth and $(n-1)$th pulses and a previous echo $=y_{n-2}+y_{n-3}+\cdots+y_1$.

(*e*) Number of quaternary mechanism echoes between nth pulse and three previous pulses $=\binom{n-1}{3}$.

(*f*) Number of quaternary mechanism echoes between the previous echoes and last three pulses $=y_{n-3}+y_{n-4}+\cdots+y_1$.

(*g*) Number of $(n-1)$-mechanism echoes between nth pulse and $(n-2)$ previous pulses $=\binom{n-1}{n-2}$.

(*h*) Number of $(n-1)$-mechanism echoes between the last pulses and previous echoes $=y_2+y_1$.

(*i*) Number of n-mechanism echoes $=\binom{n-1}{n-1}=1$.

We therefore have

$$\begin{aligned} y_n &= \binom{n-1}{1}+\binom{n-1}{2}+\cdots+\binom{n-1}{n-1} \\ &\quad +[y_{n-1}+2y_{n-2}+\cdots+ry_{n-r}+\cdots \\ &\qquad +(n-2)y_2+(n-1)y_1] \\ &= (2)^{n-1}-1+[y_{n-1}+2y_{n-2}+\cdots+ry_{n-r}+\cdots \\ &\qquad +(n-2)y_2+(n-1)y_1]. \end{aligned} \tag{3}$$

By using $y_1=0$, the formula (3) gives $y_2=1$, $y_3=4$, $y_4=13$, $y_5=39$, $y_6=112$, in agreement with the results of direct calculations with 2, 3, 4, 5 and 6 pulses, respectively. The laws of combination giving rise to the different echoes with four pulses are shown in Fig. 2, where a horizontal line is drawn to represent the mode of combination for each echo, the dots representing the parent pulses and the crosses the echoes obtained from them.

Of course it is to be remembered that the recursion formula (3) holds only under the conditions:

$$\left.\begin{aligned} &\tau_n>2\tau_{n-1}, \quad \tau_{n-1}>2\tau_{n-2}, \quad \text{etc.}; \\ &\text{or, in terms of } \delta, \\ &\delta_{n-1}>\delta_{n-2}>\delta_{n-3}>\cdots>\delta_1. \end{aligned}\right\} \tag{4}$$

If these conditions do not hold, some of the echoes

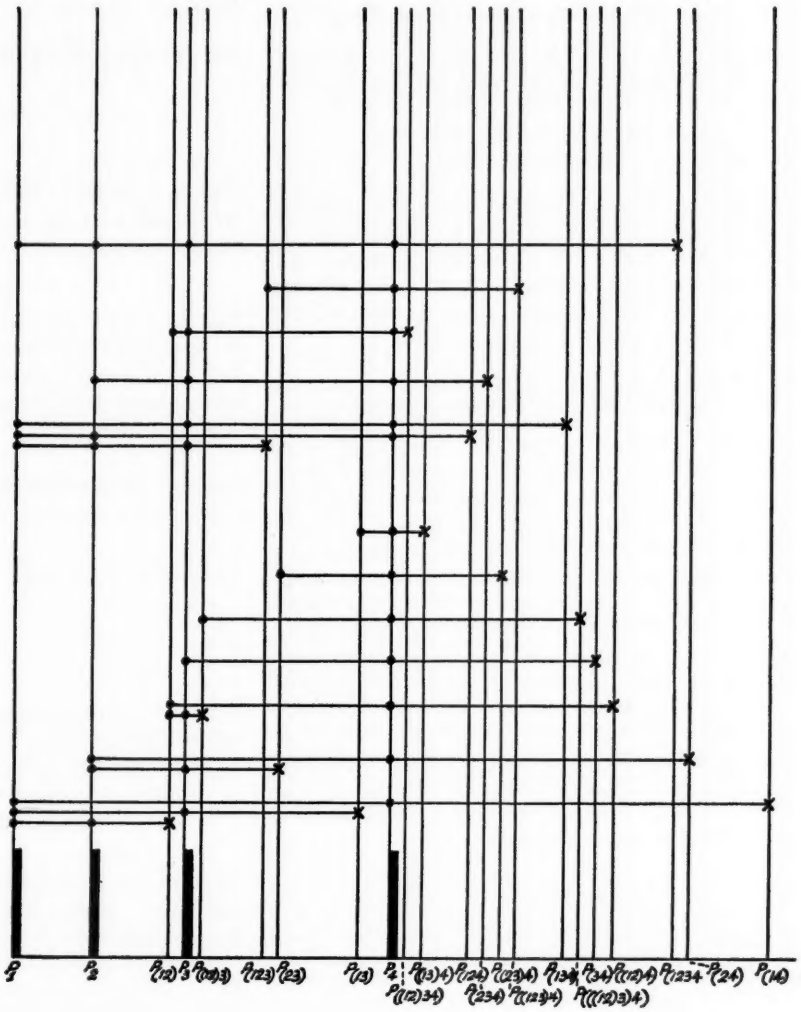


FIG. 2. Diagram illustrating the mode of formation of the echoes with 4 pulses.

predicted by (3) will not be observed. Moreover, one may enquire why we do not observe all these echoes discussed above when we apply successive pairs of pulses of increasing intervals, as in primary-echo envelope measurements.[1] The reason is that in such measurements, the pulse intervals do not evidently satisfy the conditions (4) so that some of the echoes discussed above do not occur. Further, it is to be expected that since the applied pulses now consist of pairs of pulses with fairly large gaps between successive pairs, all the echoes arising out of pulses belonging to different pairs will be heavily damped and so disappear in noise. Only the primary echoes between pulses within a pair will have significant amplitude and will be observed.

The aforementioned analysis for echoes with four pulses has an important practical application. Thus, it is difficult to set the angle of nutation $\xi=\omega_1 t_w$ at exactly $\pi/2$, because $\omega_1=\gamma H_r$, and H_r, the rf field amplitude cannot be measured accurately by a direct method. But the present investigation gives us a way to attain this end. Thus we see that the stimulated echo due to the first, second, and fourth pulses has an amplitude given by

$$P_{(124)}=-\tfrac{1}{2}\sin^3\xi\cos\xi\left(\exp-\frac{5k\tau_1^3}{3}-k\tau_1^2\tau_3\right)\times\exp\left(-\frac{2\tau_1}{T_2}-\frac{\tau_3-\tau_1}{T_1}\right). \quad (5)$$

Thus if a circuit be arranged to give four successive rf pulses with intervals satisfying (1) and also such that the diffusion and relaxation dampings are small, and then the groups of pulses be repeated at a rate faster than that necessary for persistence of vision, the full

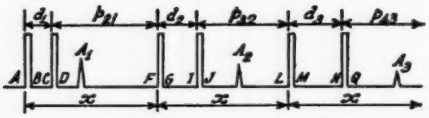


FIG. 3. Arrangements of rf pulses for Hahn primary-echo experiment.

echo pattern can be seen, from which the particular echo of interest, *viz.* $P_{(124)}$, may be chosen from its position. Now, keeping the pulse-width t_w constant, the rf voltage from the oscillator may be varied till this echo vanishes. The value of H_r when this occurs is the value necessary for making $\xi=\omega_1 t_w=\pi/2$.

IV. APPLICATIONS TO HAHN'S MEASUREMENTS ON PRIMARY AND SECONDARY ECHOES

The aforementioned analysis with four applied pulses also has interesting applications in experimental measurements on primary and stimulated echoes, with groups of two and three pulses respectively. Thus, it gives us a possible method of obtaining T_1 from primary echo measurements alone, and also a correction to measurement of T_2 from observations on primary echoes. Hahn,[1] in his experiments on the primary echo, applies paired pulses (Fig. 3) with intervals $d_1, d_2, d_3, \cdots$, respectively between members of successive pairs, the intervals between the last member of the $(n-1)$th pair and the first member of the nth pair being denoted by $p_{n,(n-1)}$. The interval between the first pulses of successive pairs is denoted by x, which is usually kept constant.

From Table I, we get the primary echo amplitude due to the first pair of pulses as

$$A_1=\sin\xi\,\sin^2(\xi/2)\,\exp\left(-\frac{2d_1}{T_2}-\frac{5kd_1^3}{3}\right).$$

The primary echo amplitude due to the second pair of pulses is given by

$$A_2=\sin\xi\,\sin^2(\xi/2)\left\{1+[W'(D)-1]\exp\left(-\frac{p_{21}}{T_1}\right)\right\}\times\exp\left(-\frac{2d_2}{T_2}-\frac{5kd_2^3}{3}\right),$$

where $W'(D)$ represents that part of $W(D)$ which is contributed by the W's of the previous instants. The term

$$C=1+\{W'(D)-1\}\exp(-p_{21}/T_1) \qquad (6)$$

gives the effect, on the primary-echo amplitude, of the memory retained by the spin system regarding the conditions established by the previous pair of pulses, and is a consequence of the fact that thermal equilibrium is not attained in the interval between the two successive pairs of pulses. Similar calculations with a third pair of pulses gives the primary-echo amplitude due to these as

$$A_3=\sin\xi\,\sin^2(\xi/2)\left\{1+[W'(J)-1]\exp\left(-\frac{p_{32}}{T_1}\right)\right\}\times\exp\left(-\frac{2d_3}{T_2}-\frac{5kd_3^3}{3}\right).$$

If we now use pulses with $\xi=\pi/2$, then the C term for the primary echo due to the nth pair of pulses will have the general form

$$[1-\exp(-p_{n,n-1}/T_1)].$$

This factor we shall call the "memory damping" factor in the primary-echo experiment. This "memory damping" effect provides us with a method of obtaining T_1 from primary-echo measurements alone. Thus, from Fig. 3, we have

$$p_{n,n-1}=x-d_{n-1}.$$

Hence the amplitude of the primary echo due to the nth pair of pulses will be given by

$$A_n=\sin\xi\,\sin^2(\xi/2)\,\exp\left(-\frac{2d_n}{T_2}-\frac{5k}{3}d_n^3\right)\times\left[1-\exp\left(-\frac{x-d_{n-1}}{T_1}\right)\right].$$

If now the d_n's be kept constant and the interval between the successive pairs of pulses be varied, then the primary-echo envelope will be given by

$$A_n\propto\left[1-\exp\left(-\frac{x-d}{T_1}\right)\right]. \qquad (7)$$

Thus, we can arrange a circuit to give us paired pulses with equal intervals between the members of each pair, but with gradually decreasing intervals x between successive pairs, starting from a value of x large compared to T_1, corresponding to attainment of thermal equilibrium between successive pairs of pulses, till x is of the order of T_1 or even smaller. By taking multiple exposures we shall then get the echo envelope giving by (7), and by taking a log-plot of this envelope we can then evaluate T_1 directly.

In Hahn's primary-echo experiment to measure T_2, on the other hand, x is kept constant and d is varied. In such a case, unless $x\gg d$, there will be a correction due to the "memory damping" factor. Thus, considering the primary echoes due to the second and third pairs of pulses, we have

$$\frac{A_2}{A_3}=\frac{1-\exp[-(x-d_1)/T_1]}{1-\exp[-(x-d_2)/T_1]}\cdot\frac{\exp(-2d_2/T_2)}{\exp(-2d_3/T_2)}.$$

If we had neglected the momemory damping factor, we would have had

$$\frac{A_2}{A_3}=\frac{\exp(-2d_2/T_2')}{\exp(-2d_3/T_2')},$$

T_2' representing the uncorrected value of T_2. Then, putting

$$d_2-d_1=d_3-d_2=\cdots=d_n-d_{n-1}=\beta$$

as employed in Hahn's experiment,[1] we get

$$\frac{1}{T_2}=\frac{1}{T_2'}-\frac{1}{2\beta}\log_e\left[1-\exp\left(-\frac{x}{T_1}\right)\frac{\beta}{T_1}\right]. \quad (8)$$

Of course, in general, the correction will be small, because, for T_1 large, β/T_1 is small and for T_1 small x/T_1 can be made large. A similar correction for "memory damping" will apply to stimulated echo measurement with groups of three pulses also.

ACKNOWLEDGMENT

The authors are very grateful to Professor M. N. Saha for his keen interest in their work. They are also thankful to Dr. A. K. Saha for guidance and encouragement.

PHYSICAL REVIEW VOLUME 98, NUMBER 2 APRIL 15, 1955

Dirac Bracket Transformations in Phase Space*

PETER G. BERGMANN AND IRWIN GOLDBERG
Department of Physics, Syracuse University, Syracuse, New York
(Received December 22, 1954)

The purpose of this paper is twofold. One is to analyze the group-theoretical significance of the Dirac bracket and to examine, in particular, the apparent ambiguities in the presence of both first-class and second-class constraints. The other is to prepare the ground for the utilization of the Dirac bracket for the quantization of generally covariant theories. It is shown that the Dirac bracket represents the commutator of infinitesimal transformations in phase space which are not canonical but form a group, in that they are the only transformations that preserve the form of all the constraints of a theory as well as the canonical form of the equations of motion. This group of transformations possesses an invariant subgroup: those transformations that correspond to coordinate transformations, gauge transformations and the like. From this subgroup, we can construct the factor group. All members of the original group have generators, but the generators of the invariant subgroup are zero. There is, then, a one-to-one correspondence between the nonvanishing generators and the members of the factor group. The Dirac bracket is uniquely defined for all dynamical variables that can serve as generators; excluded are variables that have no invariant significance (such as the divergence of the electromagnetic vector potential). If it is possible to identify all these permissible generators, then the theory can be reformulated in terms of these and will be free of constraints. It is proposed to adopt the set of generators and the Dirac brackets between them as the point of departure for the formulation of covariant quantum theories.

1. INTRODUCTION

IT is well known that the quantization of a given classical field theory is not a straightforward unique process. The replacement of classical Poisson brackets by commutators between operators cannot be carried out simultaneously for all conceivable dynamical variables without leading to internal inconsistencies. Ordinarily, the formulation of commutators is therefore restricted at first to a certain class of variables, such as the canonical coordinates of the theory,[1] all other commutators being obtained subsequently by calculations from this primary set. But a given classical field theory may be set up in terms of any set of canonically conjugate variables; the transition from Poisson brackets to operator commutators then leads to different quantum theories depending on the canonical coordinate system chosen. Hence the usual prescription is ambiguous unless we can single out a particular canonical coordinate system for the transition to quantum theory.

This situation is somewhat less ambiguous in quantum mechanics. There the rule is to carry out the quantization in Cartesian (or Lorentzian) coordinates and to go over to other coordinate systems only after quantization. This rule can be transferred to field theories that are "essentially linear." An essentially linear theory is one whose Lagrangian can be split naturally into a "free field" part and an "interaction" part, such as the Lagrangian of electrodynamics. This separation of the Lagrangian into two components, of which the former is purely quadratic in the field variables, remains preserved under linear-homogeneous variable transformations. There is, then, a set of privileged and identifiable canonical coordinate systems, any one of which may be adopted as the point of departure for quantization. The resulting theories are all equivalent to each other, except for terms that are readily identified as so-called zero-point energy terms. Such terms can be subtracted from a given Lagrangian or Hamiltonian without serious consequences. There is some doubt about the ultimate justification for such a procedure, especially in the event of strong coupling, but we shall not concern ourselves with that question here.

Generally covariant theories are not essentially linear. Moreover, the most convenient canonical coordinates of such a theory, i.e., the initial field variables

* This work was supported by the Office of Naval Research.
[1] R. E. Peierls, Proc. Roy. Soc. (London) A214, 143 (1952).

and their canonical conjugates, are not ordinarily observables in the quantum-theoretical sense. It is true that the Hilbert space of physically possible states can be imbedded in a linear vector space containing also physically meaningless states, and that all the dynamical variables of the classical theory can be identified with linear operations in this larger space. But this construction is by no means unique; it does not appear likely that with its help the unitary transformations in Hilbert space can be identified more readily with the classical canonical transformations and their generators.

Dirac[2] has opened a new approach. He has introduced into the classical theory a new type of bracket (which, with some further modifications, we shall call the Dirac bracket) that corresponds closely to mapping operations among physically meaningful states. The Dirac bracket avoids a number of mathematical pitfalls of the Poisson bracket as applied to covariant field theories. We shall give this new bracket a precise transformation-theoretical significance. In the concluding section we shall indicate the course of quantization.

Briefly, the thesis of this paper is the following. An ordinary classical theory may be described fully in terms of the structure of its group of canonical transformations. The corresponding quantum theory will have a corresponding group of unitary transformations whose commutator algebra is similar. The quantum theory can be fully characterized by a precise statement in which respects the classical group of canonical transformations and the quantum group of unitary transformations are alike. In a covariant theory there is, first of all, a group of transformations that maintains the covariance properties; this is the basic group, and it takes over the role of the canonical transformations in an ordinary theory. But this group contains a subgroup, those transformations that make up the invariance group of the theory itself (or of the class of theories). The transformations belonging to this subgroup do not map a state into a different state, but into the same state described in a different representation. (Coordinate transformations, gauge transformations and the like.) These transformations have generators that vanish.[3] They form an invariant (or normal) subgroup, hence it is possible to construct a factor group of transformations which, though preserving the invariance character of the theory, lead from one physically meaningful state to another, different one. The generators of the factor group are different from zero (except, of course, for the identity mapping); moreover, there is a one-to-one relationship between the (abstract) factor group and the generators, which may be considered as a true realization of the factor group. It is our contention that this factor group represents the appropriate point of departure for quantization. Theories possessing identical covariance properties differ from each other in the structure of the factor group. Hence the factor group is the distinguishing mark of a particular quantized theory within the set of theories possessing the required covariance properties.

This point of view will not provide a logarithm permitting the unique quantization of any given theory. But if sustained, it will make possible a more effective and unhampered analysis of the particular characteristics of a given theory while guaranteeing the desired covariance properties as a matter of course.

2. ORIGIN OF CONSTRAINTS. CANONICAL TRANSFORMATIONS

When a theory whose differential equations can be derived from a variational principle is invariant with respect to a group of transformations that involve arbitrary functions of the coordinates, including the time, then invariably the Lagrangian of such a theory is singular in the following sense: The matrix L^{AB} of the expressions that are the second mixed derivatives of the Lagrangian density with respect to the time derivatives of the field variables y_A and y_B (i.e., $\partial^2 L/\partial\dot{y}_A\partial\dot{y}_B$) is singular, and its null vectors are related to the changes in the field variables y_A under an infinitesimal transformation belonging to the invariance group.[4,5] The consequences of this singularity are manifold: the field equations are not independent of each other, but satisfy differential identities; some linear combinations of the field equations are free of second time derivatives; and though the canonical momentum densities are uniquely determined in terms of the field variables and their time derivatives, they are algebraically related to each other through what we have called the (primary) constraints.[5] If we go over to a canonical formulation of the theory, then these primary constraints occur with arbitrary coefficients in the Hamiltonian.[2,4–6] The requirement that the constraints remain satisfied in the course of time if they are satisfied at the time t_0 leads to additional constraints[3] (secondary constraints). Both the primary and the secondary constraints are what Dirac calls first-class constraints.[2] That is, the Poisson brackets between any of these constraints vanish, provided only the constraints themselves are satisfied. Moreover, the Poisson brackets between a complete set of such constraints and the Hamiltonian vanish, too, again modulo the constraints.

Aside from constraints that are related to the covariance properties of a theory we may encounter additional constraints. These arise, for instance, if in a Lagrangian certain of the time derivatives of the field variables are introduced as formally independent variables or if, for any other reasons, the time derivatives

[2] P. A. M. Dirac, Can. J. Math. **2**, 129 (1950); **3**, 1 (1951).
[3] J. L. Anderson and P. G. Bergmann, Phys. Rev. **83**, 1018 (1951).
[4] L. Rosenfeld, Ann. Physik **5**, 113 (1930); Ann. Inst. H. Poincaré **2**, 25 (1932).
[5] P. G. Bergmann, Phys. Rev. **75**, 680 (1949); P. G. Bergmann and H. M. Brunings, Revs. Modern Phys. **21**, 480 (1949).
[6] Bergmann, Penfield, Schiller, and Zatzkis, Phys. Rev. **80**, 81 (1950); R. Penfield, Phys. Rev. **84**, 737 (1951).

of certain of the field variables enter the Lagrangian linearly rather than quadratically. The most important instance of such field variables are probably those representing de Broglie waves of particles. In all cases, the derivative of the Lagrangian density with respect to one of the time derivatives is free of such time derivatives; and thus the corresponding canonical momentum density is a function of the undifferentiated field variables. These constraints generally will have nonvanishing Poisson brackets with each other,[7] though their Poisson brackets with the Hamiltonian will vanish modulo the constraints themselves. Dirac calls this type of constraint a second-class constraint.[2]

In a quantized theory that is patterned after the corresponding classical theory, the observables representing constraints must, of course, have zero expectation values. In fact, if we construct a base system of Hilbert vectors that are all physically meaningful states (i.e., in which for each base vector Ψ and for each constraint C^a the expectation value $\bar{C}^a$ vanishes), then all matrix elements of every constraint must be zero. This requirement is inconsistent with the fact that there are undoubtedly dynamical variables whose Poisson brackets with the constraints fail to vanish, which in fact are equal to 1. At least there is an inconsistency as long as we attempt to have each classical dynamical variable possess a counter-part Hermitian operator and as long as each Poisson bracket is to go over into an operator commutator. The contradiction is particularly glaring for second-class constraints: For second-class constraints it is generally possible to construct two constraints whose Poisson bracket equals 1.

Dirac[2] has constructed a new type of bracket (the Dirac bracket) which vanishes whenever one of the two factors is a second-class constraint. It differs from the ordinary Poisson bracket by terms that are bilinear in Poisson brackets involving second-class constraints. For a system with a finite number of degrees of freedom, the Dirac bracket $(A,B)^*$ is defined in terms of the Poisson brackets by the expression:

$$(A,B)^*=(A,B)+(A,C^a)(B,C^b)F_{ab}, \tag{2.1}$$

where the summations are to be extended over all second-class constraints and where the quantities F_{ab} are the elements of the matrix reciprocal to the matrix consisting of the Poisson brackets between the second-class constraints, e.g.,

$$F_{ab}(C^b,C^c)=\delta_a{}^c. \tag{2.2}$$

If a theory contains both first-class and second-class constraints, then Dirac's prescription requires that the constraints be replaced by algebraic combinations between them so that the number of second-class constraints is made as small as possible. The Dirac brackets are shown, through straightforward computation, to satisfy the Jacobi identities,

$$((A,B)^*,C)^*+((B,C)^*,A)^*+((C,A)^*,B)^*\equiv 0. \tag{2.3}$$

Hence they may be adopted as the prototypes of the quantum-theoretical commutators, instead of the usual Poisson brackets.

We shall now develop a formalism (in the c-number theory) that leads directly to the group property of the Dirac brackets and which does not require the explicit separation of the first-class and second-class constraints. In order to simplify matters, we shall demonstrate our approach on a theory with a finite number of degrees of freedom. Transition to a field theory involves no additional features.

We start out with the well-known variational principle in phase space from which the canonical equations of motion can be derived as Euler-Lagrange equations,

$$S=\int \tilde{L}dt, \quad \tilde{L}=p_k\dot{q}_k-H(q,p,t), \tag{2.4}$$

to which must be added a number of constraints, which we shall symbolize by the notation

$$C^a(q_k,p_k,t)=0. \tag{2.5}$$

When we carry out an arbitrary transformation in phase space (not necessarily a canonical transformation), then the Lagrangian $\tilde{L}$, as a function of its arguments q_k, $\dot{q}_k$, p_k, t, will change as follows:

$$\begin{aligned}\delta''\tilde{L}&=\dot{Q}-\frac{\partial\tilde{L}}{\partial q_k}\delta q_k-\frac{\partial\tilde{L}}{\partial p_k}\delta p_k-\frac{\partial\tilde{L}}{\partial\dot{q}_k}\delta\dot{q}_k\\&=\frac{d}{dt}(Q-p_k\delta q_k)+\left(\frac{\partial H}{\partial q_k}+\dot{p}_k\right)\delta q_k+\left(\frac{\partial H}{\partial p_k}-\dot{q}_k\right)\delta p_k.\end{aligned} \tag{2.6}$$

In this expression, $\dot{Q}$ is an arbitrary exact time derivative [i.e., the time derivative of an arbitrary point function in phase space, $Q(q,p,t)$], to be restricted later. In general, these additions to the Lagrangian $\tilde{L}$ will be such as to change its form, and the equations of motion in the transformed coordinate system will no longer possess the canonical form. If the canonical form is to be preserved, the change in L must have the form

$$\delta''L=-\delta''H=F(q,p,t)+\frac{\partial H}{\partial q_k}\delta q_k+\frac{\partial H}{\partial p_k}\delta p_k. \tag{2.7}$$

Comparing Eqs. (2.6) and (2.7), we have

$$\frac{d}{dt}(Q-p_k\delta q_k)+\dot{p}_k\delta q_k-\dot{q}_k\delta p_k\equiv F(q,p,t) \tag{2.8}$$

as a condition that places restrictions on both the transformation quantities δq_k, δp_k and on the choice of Q.

If we disregard the constraints, then the condition (2.8) forces us to assign to the coefficients of the time

[7] E. Newman (to be published).

derivatives of the canonical coordinates such values as to convert them into an exact time derivative. Hence we obtain the customary conditions for an infinitesimal canonical transformation:

$$\delta q_k = \partial C/\partial p_k, \quad \delta p_k = -\partial C/\partial q_k,$$
$$F = -\partial C/\partial t, \quad Q + C - p_k \delta q_k = 0. \tag{2.9}$$

The choice of a single dynamical variable, the "generating function" C, whose dependence on q, p, and t is arbitrary, determines all the other transformation quantities, including the variable Q.

3. TRANSFORMATIONS WITH CONSTRAINTS

In a self-consistent theory, the constraints are constants of the motion, at least modulo themselves. That is to say, if the constraints are all satisfied at some time t_0, then they will remain satisfied if the canonical equations of motion are obeyed. Formally, we have then:

$$\partial C^a/\partial t + (C^a, H) = \alpha^a{}_b C^b. \tag{3.1}$$

If a theory initially possesses constraints not obeying the conditions (3.1), then it will be necessary to add the left-hand sides of (3.1) to the constraints and form their time derivatives in turn, and so forth, until the iteration of this procedure produces no further constraints. This procedure may, in the case of a contrived example, also lead to unacceptable requirements on the dynamical variables. In that case the original variational principle was not a physically reasonable point of departure.

Under a transformation (2.6) or (2.9), the constraints will generally change their forms. If they do, then a peculiar difficulty arises in the subsequent quantization of such a theory. In the quantized theory, the transformation group corresponding to the infinitesimal canonical transformations is that of infinitesimal unitary transformations, which are generated by Hermitian operators in Hilbert space. These generators are then interpreted as the observables of the theory. In these presence of constraints, these two roles of the Hermitian operators may come into conflict with each other. As an observable, the constraint must vanish. As the generator of an infinitesimal unitary transformation it must not. This conflict is particularly serious in the case of "second-class" constraints. The canonical transformations generated by two suitably chosen second-class constraints are noncommutative, hence their Poisson bracket is nonzero. But, clearly, two Hermitian operators in quantum theory cannot have the eigenvalue zero with the same state vector and also possess a commutator whose expectation value is different from zero. With a first-class constraint, the situation is not much better. By definition, a first-class constraint commutes with all other constraints. Hence there must be some nonconstraint Hermitian operator with which it does not commute. The constraint, then, though itself zero for physically permissible states will generate a unitary transformation in which at least some observables change their values.

It would be much more convincing if we could formulate an operator algebra such that the constraints are not only zero in the whole Hilbert space (the latter consisting, presumably, only of permissible states) but as a result commute with every true observable. In this section we shall construct a transformation group in classical phase space that has this property.

We shall consider the group of all (infinitesimal) transformations that leave the constraints unchanged (*modulo* the constraints themselves) and which preserve the canonical form of the equations of motion. To accomplish this aim, we consider the hypersurface in phase space on which all constraints, Eqs. (2.5), are satisfied. Because of Eqs. (3.1) the motion will not lead outside this hypersurface. Likewise, the coordinate transformation in phase space must be tangential to the same hypersurface $C^a=0$ if the constraints are to remain intact. To facilitate the notation, we shall designate the q_k, p_k uniformly by ξ^ν, with the index ν running from 1 to $2n$ (n being the number of degrees of freedom). The set of constant coefficients occurring in the canonical equations of motion shall be designated by the symbol $\epsilon^{\mu\nu}$, so that the canonical equations of motion themselves take the form

$$\dot{\xi}^\mu = \epsilon^{\mu\nu} \frac{\partial H}{\partial \xi^\nu}, \quad \epsilon^{\mu\nu} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \tag{3.2}$$

$$\epsilon^{\mu\nu} + \epsilon^{\nu\mu} = 0, \quad \det|\epsilon^{\mu\nu}| = 1. \tag{3.3}$$

To identify the hypersurface itself, we shall introduce $(2n-n_0)$ parameters x^n, n_0 being the total number of constraints C^a. In other words, the hypersurface $C^a=0$ will be characterized by the ξ^ν being specified functions of the x^n. Whenever we find this convenient, we may extend the definition of the parameters in phase space, so that the $2n$ functions C^a, x^n form a coordinate system (albeit a noncanonical coordinate system) throughout phase space.

Equation (3.1) specifies that the equations of motion will not lead a representative point on our hypersurface off it. Likewise, we shall consider only such (infinitesimal) coordinate transformations in phase space that map the hypersurface on itself, which is another way of saying that the constraints do not change the form of their dependence on the canonical coordinates ξ^ν. These two statements may be expressed concisely by the two equations:

$$\dot{\xi}^\nu = \frac{\partial \xi^\nu}{\partial x^n} \dot{x}^n, \quad \delta\xi^\nu = \frac{\partial \xi^\nu}{\partial x^n} \delta x^n. \tag{3.4}$$

These conditions are equivalent to the two requirements that

$$\dot{C}^a \equiv \frac{\partial C^a}{\partial \xi^\nu} \dot{\xi}^\nu = 0, \quad \delta C^a \equiv \frac{\partial C^a}{\partial \xi^\nu} \delta\xi^\nu = 0. \tag{3.5}$$

Returning to Eq. (2.8), we find that the canonical equations will be reproduced if

$$\epsilon_{\mu\nu}\dot{\xi}^{\mu}\delta\xi^{\nu}+F=\epsilon_{\mu\nu}\frac{\partial\xi^{\mu}}{\partial x^{m}}\frac{\partial\xi^{\nu}}{\partial x^{n}}\dot{x}^{m}\delta x^{n}+F=\frac{d\Gamma}{dt}$$
$$=\frac{\partial\Gamma}{\partial x^{m}}\dot{x}^{m}+\frac{\partial\Gamma}{\partial t}. \tag{3.6}$$

We are thus led to the requirements:

$$\epsilon_{mn}\delta x^{n}=\partial\Gamma/\partial x^{m},\quad F=\partial\Gamma/\partial t=\delta H,$$
$$Q+\Gamma-p_{k}\delta q_{k}=0. \tag{3.7}$$

In the first of these relations, the coefficients ϵ_{mn} are short for

$$\epsilon_{mn}=\frac{\partial\xi^{\mu}}{\partial x^{m}}\frac{\partial\xi^{\nu}}{\partial x^{n}}\epsilon_{\mu\nu}\equiv\xi^{\mu}{}_{|m}\xi^{\nu}{}_{|n}\epsilon_{\mu\nu},$$
$$\epsilon_{\mu\nu}=\begin{pmatrix}0 & -I\\ I & 0\end{pmatrix},\quad \epsilon_{\mu\rho}\epsilon^{\rho\nu}=\delta_{\mu}{}^{\nu} \tag{3.8}$$

and are antis mmetric in their subscripts m,n. (In what follows we shall designate differentiation with respect to the canonical coordinates by the comma symbol, differentiation with respect to a constraint by the vertical bar, followed by a subscript a, b, $\cdots$, and differentiation with respect to a parameter also by the vertical var, followed by a subscript m, n, ...) All these relations are to be considered valid only on the hypersurface $C^{a}=0$. Hence the dependence of the new generating function Γ on the C^{a}, does not affect the situation; we may add to a generating function any expression that vanishes together with the constraints, without affecting the resulting transformation. Equations (3.8) are also manifestly invariant with respect to any transformation of the parameters such that the new parameters x' are algebraically independent functions of the x^{n} and C^{a} and the new constraints C'^{a} algebraically independent functions of the original constraints C^{a} and the parameters x^{m}, as long as $C^{a}=0$ and $C'^{a}=0$ imply each other. We conclude that any general algebraic property of the new coefficients ϵ_{mn} is independent of these choices.

Depending on the constraints C^{a}, the matrix ϵ_{mn} may be regular or singular. We shall find that ϵ_{mn} will be regular if all the constraints are second-class constraints, and singular in the presence of first-class constraints. Moreover, the number of algebraically independent second-class constraints (the minimum number that can be achieved) equals the rank of the matrix ϵ_{mn}. To prove these assertions we shall assume that there are S null vectors, designated by the symbol $U^{n}{}_{(s)}$,

$$\epsilon_{mn}U^{n}{}_{(s)}\equiv\xi^{\mu}{}_{|m}\epsilon_{\mu\nu}\xi^{\nu}{}_{|n}U^{n}{}_{(s)}=0. \tag{3.9}$$

We multiply Eq. (3.9) by $x^{m}{}_{,\rho}$ and obtain:

$$0=x^{m}{}_{,\rho}\xi^{\mu}{}_{|m}\xi^{\nu}{}_{|n}\epsilon_{\mu\nu}U^{n}{}_{(s)}$$
$$=(\delta_{\rho}{}^{\mu}-C^{a}{}_{,\rho}\xi^{\mu}{}_{|a})\epsilon_{\mu\nu}\xi^{\nu}{}_{|n}U^{n}{}_{(s)},$$
$$\epsilon_{\rho\nu}\xi^{\nu}{}_{|n}U^{n}{}_{(s)}=C^{a}{}_{,\rho}\xi^{\mu}{}_{|a}\epsilon_{\mu\nu}\xi^{\nu}{}_{|n}U^{n}{}_{(s)}\equiv C^{a}{}_{,\rho}\epsilon_{an}U^{n}{}_{(s)},$$
$$\epsilon_{an}=\xi^{\mu}{}_{|a}\xi^{\nu}{}_{|n}\epsilon_{\mu\nu}. \tag{3.10}$$

In view of the circumstance that $\epsilon_{\mu\nu}$ is a regular matrix, its reciprocal $\epsilon^{\mu\nu}$ exists (the elements are all 0, 1, and -1),

$$\epsilon_{\mu\rho}\epsilon^{\rho\mu}=\delta_{\rho}{}^{\mu} \tag{3.11}$$

and we can multiply Eq. (3.10) by it:

$$\xi^{\nu}{}_{|n}U^{n}{}_{(s)}=\epsilon^{\nu\sigma}C^{a}{}_{,\sigma}\xi^{\mu}{}_{|a}\epsilon_{\mu\rho}\xi^{\rho}{}_{|n}U^{n}{}_{(s)},$$
$$=\epsilon^{\nu\rho}C^{a}{}_{,\rho}\epsilon_{an}U^{n}{}_{(s)}. \tag{3.12}$$

We shall multiply this equation by $C^{b}{}_{,\nu}$, noting, incidentally, that

$$A,{}_{\rho}B,{}_{\sigma}\epsilon^{\rho\sigma}=(A,B) \tag{3.13}$$

is the usual Poisson bracket between the two arbitrary functions A and B. We obtain:

$$0=(C^{a},C^{b})\epsilon_{an}U^{n}{}_{(s)}$$
$$\equiv(C^{a},C^{b})U_{(s)a},$$
$$U_{(s)a}\equiv\epsilon_{an}U^{n}{}_{(s)}. \tag{3.14}$$

The matrix of Poisson brackets between the constraints is the one used by Dirac to determine the maximum number of first-class constraints that can be obtained by algebraic combination. We must still show that the S vectors $U_{(s)a}$ do not vanish. We multiply Eq. (3.12) by $x^{m}{}_{,\nu}$:

$$U^{m}{}_{(s)}=(x^{m},C^{a})U_{(s)a} \tag{3.15}$$

because of the definition, last Eq. (3.14).

We shall now return to the first Eq. (3.7). If the matrix ϵ_{mn} is regular, then we may introduce the reciprocal matrix ϵ^{mn}, multiply through, and obtain an explicit expression for the transformation quantities δx^{n}:

$$\delta x^{m}=\epsilon^{mn}\Gamma_{|n}. \tag{3.16}$$

The final expression for the transformation law is:

$$\delta\xi^{\rho}=\epsilon^{*\rho\sigma}\Gamma_{,\sigma},\quad \epsilon^{*\rho\sigma}=\xi^{\rho}{}_{|m}\epsilon^{mn}\xi^{\sigma}{}_{|n}. \tag{3.17}$$

This transformation law is not identical with that propounded in Eq. (2.9), because the new matrix $\epsilon^{*\rho\sigma}$ is not identical with the matrix $\epsilon^{\rho\sigma}$. We have, for instance,

$$C^{a}{}_{,\rho}\epsilon^{*\rho\sigma}=0. \tag{3.18}$$

If we form the new bracket (the "Dirac bracket")

$$\{A,B\}=A_{,\rho}B_{,\sigma}\epsilon^{*\rho\sigma}\equiv\delta_B A \\ =A_{|m}B_{|m}\epsilon^{mn}, \tag{3.19}$$

then we find that the Dirac bracket of a constraint with any other dynamical variable vanishes. It follows that if two dynamical variables Γ and $\overline{\Gamma}$ differ from each other by an expression that vanishes when the constraints are satisfied, then they generate the same transformation.

The group-theoretical commutator of two transformations (3.17) must be a transformation of the same type, because Eq. (3.17) represents the most general point-to-point transformation in phase space that leaves a certain set of (essentially second-class) constraints invariant. Let us examine the generator of the commutator. We have, as a general rule, for the commutator of two point-to-point transformations in any kind of space the expression

$$\delta x^\mu=(\delta_1 x^\mu)_{,\rho}\delta_2 x^\rho-(\delta_2 x^\mu)_{,\rho}\delta_1 x^\rho. \tag{3.20}$$

If the expression for the transformation law (3.19) is substituted into the commutator law (3.20), then a brief calculation shows that the commutator transformation is in fact generated by the Dirac bracket of the two generators of the commuted transformations. In the (otherwise straightforward) calculation it is important that the matrix ϵ_{mn} satisfies the cyclic commutation law

$$\epsilon_{ij|k}+\epsilon_{ki|j}+\epsilon_{jk|i}=0. \tag{3.21}$$

Because of the group-theoretical significance of the Dirac bracket, the Jacobi relationship,

$$\{\{A,B\},C\}+\{\{B,C\},A\}+\{\{C,A\},B\}=0, \tag{3.22}$$

can be inferred without any additional calculations.

Next we come to the more difficult case of a singular matrix. Now we have S null vectors satisfying Eq. (3.15). If we multiply the first set of Eqs. (3.7) by one of these null vectors, we obtain S restrictions on the generating function,

$$\Gamma_{|n}U^n{}_{(s)}=\Gamma_{,\rho}\xi^\rho{}_{|n}U^n{}_{(s)}=0. \tag{3.23}$$

But this restriction on the generating function is matched by a lack of determination of the transformation itself by the generating function. Given a generating function that satisfies the condition (3.23), we find in accordance with Eq. (3.7) that the transformation quantities δx^n are thereby determined only up to a linear combination of null vectors. In particular it is conceivable that we have a nonvanishing transformation with a zero generating function. We shall now proceed to determine the generating function of a commutator transformation. Given two transformations identified by the indices 1 and 2, respectively, we have, according to Eqs. (3.7) and (3.20), for the generating function of the commutator the expression:

$$\begin{aligned}\Gamma_{|n}&=\epsilon_{ni}\delta x^i=\epsilon_{ni}(\delta_1 x^i)_{|j}\delta_2 x^j-\epsilon_{nj}(\delta_2 x^j)_{|i}\delta_1 x^i \\ &=\Gamma_{1|nj}\delta_2 x^j-\Gamma_{2|ni}\delta_1 x^i-(\epsilon_{ni|j}-\epsilon_{nj|i})\delta_1 x^i\delta_2 x^j \\ &=\Gamma_{1|nj}\delta_2 x^j-\Gamma_{2|ni}\delta_1 x^i+\epsilon_{ij|n}\delta_1 x^i\delta_2 x^j \\ &=(\Gamma_{1|j}\delta_2 x^j-\Gamma_{2|i}\delta_1 x^i+\epsilon_{ij}\delta_1 x^i\delta_2 x^j)_{|n},\end{aligned} \tag{3.24}$$

$$\Gamma=\Gamma_{1|j}\delta_2 x^j-\Gamma_{2|i}\delta_1 x^i+\epsilon_{ij}\delta_1 x^i\delta_2 x^j.$$

Because of Eqs. (3.15) and (3.23), the value of Γ is not affected if we add to the two commuted transformation laws arbitrary combinations of the S null vectors $U^n{}_{(s)}$. Hence Γ is uniquely determined by the two generators Γ_1 and Γ_2. We shall call the expression on the right in the last line of Eq. (3.24) the (generalized) Dirac bracket. It satisfies all the properties that a commutator must satisfy, including the Jacobi relationship. It is properly defined only for variables Γ_1 and Γ_2 which in their own right satisfy the condition (3.23).

Equation (3.24) leads in particular to the result that if one of the two generating functions Γ_1 and Γ_2 vanishes, if, in other words, one of the two transformation laws is but a linear combination of null vectors, then the commutator will also have a vanishing generating function. Group-theoretically this result means that the set of transformations whose generating function vanishes is not only a subgroup, but an invariant subgroup of the group (3.7), and that there exists a quotient group. Fortunately, we already possess a realization of the quotient group: Each coset within the big group (and hence each member of the quotient group) is characterized by exactly one dynamical variable satisfying the condition (3.23). Their commutators are the expressions in the last line of Eq. (3.24). Henceforth, we shall write:

$$\{A,B\}=A_{|m}\delta_B x^m-B_{|m}\delta_A x^m+\epsilon_{mn}\delta_A x^m\delta_B x^n \\ \equiv A_{|m}\delta_B x^m\equiv-B_{|m}\delta_A x^m\equiv\epsilon_{mn}\delta_A x^m\delta_B x^n. \tag{3.25}$$

The quantized theory must incorporate a corresponding operator algebra. It need not contain operators whose classical analogs fail to satisfy the condition (3.23).

4. CONNECTION WITH DIRAC'S ORIGINAL BRACKETS

The formalism developed in Sec. 3 relieves us of the necessity of separating the constraints completely into first-class and second-class constraints by a series of algebraic combinations that makes the number of second-class constraints as small as possible. Nevertheless, the existence of these two types of constraint may be inferred from the properties of the matrix ϵ_{mn}. If this matrix is regular, then it is possible to introduce a set of parameters x^m such that all the parameters have vanishing Poisson brackets with all the constraints, and the latter will all be second-class constraints. The existence of null vectors, on the other hand, presages

the possibility of constructing first-class constraints, i.e., expressions of the form $U_{(s)a}C^a$. If it were possible to convert all the remaining second-class constraints into pairs of canonically conjugate coordinates, then our rules would provide that generators should be functions of those canonical coordinates that are neither constraints nor canonically conjugate to first-class constraints, and that brackets are to be formed with the help of these remaining coordinates only.

The chief difference between our approach and Dirac's is that Dirac permits both the first-class constraints and the coordinates conjugate to them to form generators and to be used for the formation of brackets. In practice, it would be extremely difficult to carry out the separation and classification of constraints along the simple lines indicated above, and both Dirac's prescriptions and ours endeavor to define brackets with a minimum of necessary rearrangements. Ostensibly, the two sets of rules of procedure are sufficiently different so that we shall show in this section that insofar as second-class constraints are concerned they are equivalent. Dirac's prescription is summarized in Eq. (2.1), ours in Eq. (3.25). In order to show the equivalence, we shall first remark that the Dirac bracket is uniquely determined by the two requirements (1) that the difference between the Dirac and Poisson brackets be bilinear in the Poisson brackets of A and B with the second-class constraints, and (2) that the Dirac bracket of any A with any second-class constraint C^a vanish. Since our bracket (3.25) assuredly satisfies the second requirement, it remains to prove that it also obeys the first one.

To this end we shall first put Eq. (3.25) into a more convenient form. If we multiply Eq. (3.7) by η^{lm}, the "quasi-inverse" to ϵ_{mn}, we are led to the expression

$$\delta x^n = U^n{}_{(s)} V^{(s)}{}_m \delta x^m - \eta^{mn}\Gamma_{|m}. \tag{4.1}$$

The quasi-inverse[8] η^{lm} to the singular matrix ϵ_{mn} is defined by

$$\eta^{lm}\epsilon_{mn}\eta^{nk} = \eta^{lk}, \quad \epsilon_{km}\eta^{mn}\epsilon_{nl} = \epsilon_{kl}, \tag{4.2}$$

hence

$$\eta^{lm}\epsilon_{mn} = \delta_n{}^l - U^l{}_{(s)} V_n{}^{(s)}, \tag{4.3}$$

where $U^l{}_{(s)}$ and $V_m{}^{(s)}$ are the null vectors of ϵ_{mn} and η^{lm}, respectively. Then with the aid of Eq. (3.23), Eq. (3.25) can be written as

$$\{A,B\} = +A_{|m}B_{|n}\eta^{mn}. \tag{4.4}$$

We shall now obtain the explicit form of η^{mn} in order to examine the relationship between the two bracket expressions (2.1) and (4.4). By straightforward calculation we find, first, that

$$\begin{aligned}(x^m,x^s)\epsilon_{sn} &= \delta_n{}^m - (x^m,C^a)\epsilon_{an},\\ (C^a,x^s)\epsilon_{sn} &= -(C^a,C^b)\epsilon_{bn},\end{aligned} \tag{4.5}$$

where ϵ_{an} is short for

$$\epsilon_{an} = \xi^\mu{}_{|a}\epsilon_{\mu\nu}\xi^\nu{}_{|n}. \tag{4.6}$$

Next we introduce the quantity F_{ab}, which occurs in Eq. (2.1) and which was originally understood as the reciprocal of (C^a,C^b) (because we had assumed that the latter matrix was nonsingular), as the quasi-inverse[8] of (C^a,C^b). Dirac's prescription leads, in the presence of first-class constraints, just to that quasi-inverse. We have then

$$(x^m,x^n)^* = (x^m,x^n) - (x^m,C^a)F_{ab}(C^b,x^n). \tag{4.7}$$

With the help of Eqs. (4.5), we can easily verify that

$$(x^m,x^s)^*\epsilon_{sn} = \delta_n{}^m - (x^m,C^a)U_{(s)a}V^{(s)b}\epsilon_{bn} \tag{4.8}$$

and

$$\epsilon_{mr}(x^r,x^s)^*\epsilon_{sn} = \epsilon_{mn}. \tag{4.9}$$

It follows that η^{mn} and $(x^m,x^n)^*$ can differ at most by an expression that is bilinear in the null vectors $U^m{}_{(s)}$ of the matrix ϵ_{mn}. This is indeed the case. A somewhat lengthy calculation, which we shall not reproduce here, yields for the quasi-inverse η^{mn} the expression:

$$\begin{aligned}\eta^{mn} &= (x^m,x^n)^* + U^m{}_{(r)}\eta^{(r)(s)}U^n{}_{(s)},\\ \eta^{(r)(s)} &= -V^{(r)a}\xi^\alpha{}_{|a}\epsilon_{\alpha\beta}\xi^\beta{}_{|b}V^{(s)b}.\end{aligned} \tag{4.10}$$

According to Eq. (4.4), we have, therefore, for two generators that satisfy the conditions (3.23)

$$\{A,B\} = A_{|m}(x^m,x^n)^*B_{|n}. \tag{4.11}$$

On the other hand, Eq. (2.1) yields

$$\begin{aligned}(A,B)^* = A_{|m}(x^m,x^n)^*B_{|n} + A_{|m}U^m{}_{(s)}V^{(s)b}B_{|b}\\ - A_{|a}V^{(s)a}U^n{}_{(s)}B_{|n}\end{aligned} \tag{4.12}$$

whether or not the conditions (3.23) are obeyed. If they are, then the two brackets (4.11) and (4.12) are equal.

5. CONCLUSION

To complete the present theory, we are planning to examine the group of transformations in configuration space most nearly corresponding to the phase space group developed here. It is to be assumed that so-called coordinate conditions in covariant theories (or gauge conditions in electrodynamics) will not affect the factor group, but only the invariant subgroup. To this extent, it must be possible to show the equivalence of quantization with and without such conditions.

As for quantization, it appears now that the development of the factor group is a task that should be completed prior to quantization. To give an instance, the two constraints in electrodynamics are (in a system of units in which $c=1$ and which is rationalized):

$$\pi^4 = 0, \quad \mathrm{div}\pi + \sigma = 0, \tag{5.1}$$

[8] P. G. Bergmann and R. Schiller, Phys. Rev. 89, 4 (1953).

a possible choice for the parameters is, therefore:

$$\Phi,\mathbf{A},\pi^*(\mathbf{r})=\pi(\mathbf{r})$$

$$+\int_{r'} d^3r'(|\mathbf{r}-\mathbf{r}'|)^{-1}\,\text{grad}\,\text{div}\pi(\mathbf{r}'). \quad (5.2)$$

This system, however, does not represent a point of departure for the construction of generators. There are two null vectors at each space point; Eq. (3.23) implies that the generators must be independent of Φ and of div$\mathbf{A}$. It remains, then, that generators may be constructed as (otherwise arbitrary) functions of curl$\mathbf{A}$ and of curlπ. With a given distribution of charge and current density, these are, indeed, the only variables that may be chosen at will throughout space at some particular initial time t_0; thence Maxwell's equations determine uniquely the values of all the field components. Once the Hamiltonian is formulated as a function of this set of variables, then the (classical or quantum) theory is fully determined.

This new procedure, then, avoids one of the most difficult, and as far as we are aware, hitherto unsolved problems in the quantization of generally covariant field theories: to render the Hamiltonian and the constraints compatible to each other by a suitable choice of factors. This problem is avoided and replaced by another one. It is to replace the original field variables and their canonically conjugate momentum densities, which satisfy some constraints (eight in the general theory of relativity) by a smaller number of variables, the "parameters" x^m, further diminished by the conditions (3.23). In Maxwell theory, for instance, the number of canonical variables per space point is eight, the number of constraints two, and the number of null vectors also two, so that the stripped theory contains only the equivalent of four canonical variables per space point. In the general theory of relativity we start with twenty canonical variables; the number of constraints is eight, and the number of null vectors also eight; we are again left with the equivalent of four observables per space point.

It is fairly easy to find a number of parameters in any given theory, but the determination of the null vectors is undoubtedly very difficult. It is, however, a task to be carried out in the framework of the c-number theory, and, therefore, probably less difficult than the determination of a set of compatible Hamiltonian and constraint expressions in the q-number theory. Once the observables have been determined, then presumably any Hermitian combination, which in the limit $\hbar\to 0$ goes over into the classical Hamiltonian represents a formally possible quantum theory. Then it will be necessary to develop new physical (rather than formal) criteria for the appropriate factor sequence. There are indications that the Hamiltonian of the general theory of relativity may vanish and that all the observables are constants of the motion.

Letters to the Editor

PUBLICATION of brief reports of important discoveries in physics may be secured by addressing them to this department. The closing date for this department is five weeks prior to the date of issue. No proof will be sent to the authors. The Board of Editors does not hold itself responsible for the opinions expressed by the correspondents. Communications should not exceed 600 words in length and should be submitted in duplicate.

Measurement and Interpretation of Conductance of *P*-Type Inversion Layers on Germanium*

G. A. DEMARS, H. STATZ, AND L. DAVIS, JR.
Raytheon Manufacturing Company, Waltham, Massachusetts
(Received January 10, 1955)

INVERSION layers in both *n*- and *p*-type germanium have previously been observed,[1–5] but quantitative measurements reported only for the *n*-type inversion layers.[1,2] Results of similar measurements on *p*-type inversion layers are described here and explained in terms of a constant Fermi level at the surface (a high density of surface states).

It is believed that the theory of oxidation,[6] as developed for metals, describes correctly the situation of oxidized germanium surfaces. In this theory, a strong electric field at the germanium-germanium oxide interface is expected, arising from the negative surface charge. On *n*-type germanium this field does not end at the surface but it can penetrate a considerable distance into the germanium, thus producing a space charge layer or even an inversion layer (Fig. 1). For a quantitative interpretation, the surface states which lie at the germanium-germanium oxide interface[7] have to be taken into account. These same states are considered to be responsible for surface recombination.

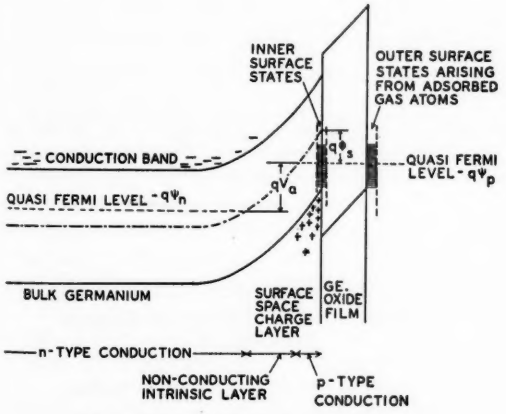


FIG. 1. Energy band structure at surface.

P-type inversion layers can be produced in many ways.[4] In this laboratory, the germanium was etched in CP4, washed, and dried. Exposure of the sample to a mixture of wet O_2 and O_3, followed by a dry mixture of O_2 and O_3 created a stable *p*-type inversion layer. Dry N_2 was also used, but the inversion layer so produced decayed gradually over a period of several hours. Active oxygen seemed to be essential in obtaining an inversion layer.

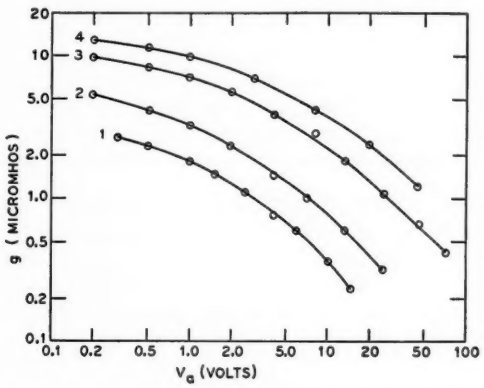


FIG. 2. Steady-state channel conductance *versus* bias voltage.

In Fig. 2, the measured conductance values $g(V_a)$ of four different inversion layers on 8 ohm-cm *n*-type material are given. $p-n-p$ grown junction germanium bars, with 0.4-cm *n*-type regions, were used. The four runs were made for surfaces which had been exposed for different times to wet oxygen and ozone. The longer the exposure, the higher the conductance of the inversion layer. All conductance readings are normalized and the numbers given represent the conductance of a unit square of *p*-type skin. All values are steady state. For each bias voltage, it was found that it took a time ranging from a few seconds for the lowest curve to a few minutes for the highest curve of Fig. 2 to establish equilibrium. The explanation for the slow adjustment will be given in a following letter.[8] For the interpretation of the conductivity values of the channel, the position of the quasi-Fermi level is calculated with respect to the middle of the band at the surface, i.e., ϕ_s (Fig. 1). The total positive space charge Q inside the semiconductor, which results from ionized donor atoms and holes is also computed. This space charge must be compensated by electrons in surface states which lie at the germanium-germanium oxide interface and at the germanium oxide surface. The same quasi-Fermi level ψ_p which describes the occupancy of the states in the valence band just inside the germanium will also describe the occupancy of the two types of surface states. Thus, the relation between ϕ_s and Q will give

TABLE I. Position of the quasi-Fermi level at the surface and the total charge in the surface states for various bias voltages.x (Steady-state case.)

	Run 1		Run 2		Run 3		Run 4	
V_a	ϕ_s (volts)	No. electronic charges/cm² ×10⁻¹⁰	ϕ_s (volts)	No. electronic charges/cm² ×10⁻¹⁰	ϕ_s (volts)	No. electronic charges/cm² ×10⁻¹⁰	ϕ_s (volts)	No. electronic charges/cm² ×10⁻¹⁰
0.2	0.110	4.97	0.134	6.50	0.158	9.42	0.169	11.4
0.5	0.117	6.26	0.135	7.37	0.160	10.2	0.171	12.3
1	0.119	7.75	0.137	8.70	0.162	11.5	0.173	13.6
3	0.121	11.7	0.138	12.4	0.164	14.8	0.178	16.9
5	0.122	14.7	0.138	15.2	0.165	17.2	0.180	19.4
7	0.122	17.1	0.138	17.5	0.167	19.5	0.182	21.8
10	0.120	20.2	0.138	20.4	0.167	22.3	0.183	24.3
12	0.120	21.9	0.137	22.3	0.167	24.1	0.183	25.9
14	0.118	23.6	0.137	23.9	0.167	25.5	0.183	27.3
15	0.117	24.4	0.137	24.7	0.167	26.3	0.183	28.0
16			0.137	25.5	0.167	27.0	0.183	28.7
18			0.137	26.9	0.168	28.4	0.185	30.2
20			0.137	28.4	0.168	29.8	0.184	31.4
25			0.136	31.6	0.168	32.7	0.184	34.3
45					0.167	42.9	0.183	44.0
70					0.167	52.9		

information about the total density of surface states (i.e., the sum of the densities of the two states).

The computation of ϕ_s and Q has been carried out by using the mobility for holes in inversion layers as calculated by Schrieffer.[9] This analysis does not make use of the approximations described by Brown[1] or Kingston.[10] The results for the four runs are given in Table I.

It is seen that ϕ_s is essentially constant for each run. It is uncertain whether any significance can be attributed to the deviation of ϕ_s from a constant. The total charge in the surface traps continues to increase with increasing bias voltage. If the slight increase of ϕ_s for voltages approximately below 10 volts is real, then this represents a serious difficulty. From Fig. 1, it may be seen that an increase in ϕ_s means a decrease in the occupancy of the surface states. On the other hand, the requirement of total neutrality gives an increasing charge in the surface states. This difficulty can be avoided by assuming some specular reflection of holes at the surface in Schrieffer's calculation.[9] Another possible source of error is Schrieffer's simplifying but incorrect assumption of spherical energy surfaces. In addition, the base resistivity used in the calculation may have been incorrect by as much as 20 percent.

In every case, it may be said that the Fermi level is essentially constant and the change in charge very considerable, so that the total density of surface states will be high. It is not possible to give a numerical result for the density of states, because the Fermi level shifts by an amount that is smaller than the probable error of this determination. However, from the measurements described in the accompanying letter,[8] it may be concluded that this high total density of states is due only to surface states outside the germanium oxide. The charge in the states at the germanium-germanium oxide surface does not change by such a large amount for so little change in the Fermi level.

Kingston[10] has reported that the data for channels on p-type germanium can also be understood by assuming a constant value of ϕ_s. It is interesting to note that our curves, which deviate greatly from the $1/V_a$ relationship measured by Kingston, still give approximately a constant ϕ_s value, indicating that Kingston's approximations are not applicable in our case.

* This work has been supported in part by the U. S. Signal Corps.

[1] W. L. Brown, Phys. Rev. **91**, 518 (1953).
[2] R. H. Kingston, Phys. Rev. **93**, 346 (1954).
[3] S. R. Morrison, J. Phys. Chem. **57**, 860 (1953).
[4] H. Christensen, Proc. Inst. Radio Engrs. **42**, 1317 (1954).
[5] A. L. McWhorter and R. H. Kingston, Proc. Inst. Radio Engrs. **42**, 1376 (1954).
[6] N. Cabrera and N. F. Mott, Repts. Progr. Phys. **12**, 163 (1949).
[7] W. H. Brattain and J. Bardeen, Bell System Tech. J. **32**, 1 (1953).
[8] Statz, Davis, and deMars, following Letter [Phys. Rev. **98**, 540 (1955)].
[9] J. R. Schrieffer, Phys. Rev. **97**, 641 (1955). The authors thank Mr. Schrieffer for a preprint of this article.
[10] R. H. Kingston (to be published). The authors are grateful for the privilege of seeing a preprint of Dr. Kingston's paper.

Structure of Surface States at the Germanium-Germanium Oxide Interface*

H. STATZ, L. DAVIS, JR., AND G. A. DEMARS
Raytheon Manufacturing Company, Waltham, Massachusetts
(Received January 10, 1955)

IN the accompanying Letter,[1] the conductance of p-type inversion layers on 8 ohm-cm n-type germanium under steady-state conditions has been discussed. In this Letter, investigations of the nonsteady-state conductance and its interpretation in terms of surface states at the germanium-germanium oxide interface are reported.

The time dependence of the conductance g is shown schematically (Fig. 1) when the bias voltage was switched from a value V_1 to a value V_2. The time-dependent portion of the conductance has the form $\exp(-t/\tau)$, except when the voltage steps $|V_1-V_2|$ were large compared to V_1.

The same $p-n-p$ grown junction germanium samples (8 ohm cm n-type) with the same surface treatment as in reference 1 are used. The longer the surface is exposed to wet O_2 and O_3, the higher the inversion layer conductance and the longer the time constant τ. The time constant varied from a few seconds to a few minutes and also increased with bias voltage for a given inversion layer.

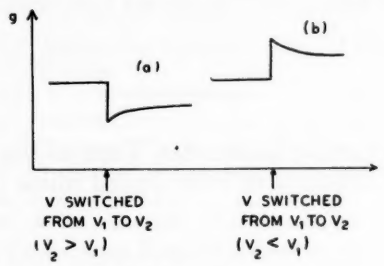


FIG. 1. Time dependence of conductance when bias voltage is switched.

As may be seen in Fig. 1, the conductance after switching to a higher bias voltage was lower than its steady-state value. The inversion layer contained fewer holes and therefore the total charge in the surface was smaller than under steady-state conditions. If the voltage was switched to a lower value, the opposite was true. These observations can be understood from the fact that the charge in the surface states under steady-state conditions continuously increased with increasing bias voltage. For example, when the bias voltage was switched to higher values, there must have been a transfer of electrons from the germanium to surface states at the germanium-germanium oxide interface and germanium oxide surface (Fig. 1, reference 1). From the measurements, it was concluded that the transfer of electrons was at least partially a slow process. The long time constant can be connected only with charge transfer to states lying at the germanium oxide surface. The inner states at the germanium-germanium oxide interface which were probably responsible for the surface recombination[2] adjusted to a new quasi-Fermi level in a time which was presumably of the order of the lifetime of the minority carriers in the sample. With the equipment used, changes of such a short duration could not be resolved. It must therefore be assumed that just after the voltage V_2 was applied, the conductance reading corresponded to a situation in which the charge in the outer surface states was still unchanged, but the charge in the inner surface states was already in equilibrium with the bulk of the germanium. The good fit of the data so obtained, as seen later in Fig. 3, can be considered as strong evidence for the validity of this assumption.

The widely different time constants for the two types of surface states permit one to determine the density of the inner states separately. The measurements were made as follows: The bias voltage V_a was set to some particular value for a long enough time so that the conductance of the inversion layer had reached its steady state value, then switched to some higher value. The conductance, just after switching, was observed, extrapolating back to zero time, if necessary. The original bias was then restored and after a long enough time, the bias voltage was again switched to some other value, the conductance, just after switching, being observed again. Many points (g,V_a) were determined in this way until a smooth curve could be drawn through them. Two typical curves for two different surfaces are shown in Fig. 2. For high enough voltages, zero conductance, i.e., pinch-off, can always be obtained. By switching to lower voltages, the range of the curves could be increased.

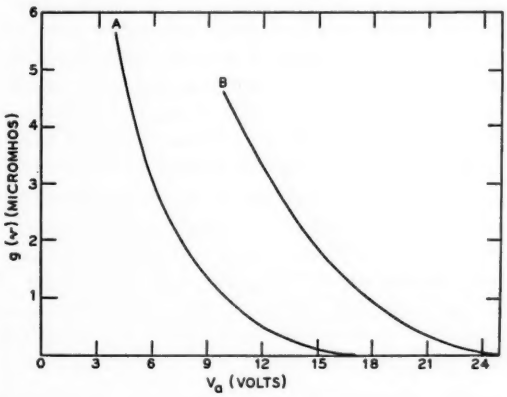


FIG. 2. Nonsteady-state channel conductance *versus* bias voltage.

The curves of Fig. 2 were analyzed in the same way as steady-state curves in reference 1. The total charge in the surface states and the position of the Fermi level with respect to the middle of the energy gap at the surface (ϕ_s) were computed. However, because of the nature of the experiment, the charge outside the germanium oxide was constant. All the variation in charge along such a curve will be attributed to the inner surface states only.

In Table I, ϕ_s and the total charge are given for the two curves of Fig. 2. According to Brattain and Bardeen, the surface recombination centers lie far away from the middle of the band. Therefore, the charge at pinch-off, i.e., $\phi_s \approx 0$, was taken as the reference point. In Fig. 3, this charge difference is plotted *versus* ϕ_s for the two cases. The result can be easily interpreted.

The solid curves are theoretical (Fermi distribution) assuming, in both cases, a single level 0.155 ev below

TABLE I. Position of the quasi-Fermi level at the surface and the total charge in the surface states for various bias voltages (nonsteady-state case).

Run A			Run B		
V_a	ϕ_s (volts)	No. electronic charges/cm² $\times 10^{-11}$	V_a	ϕ_s (volts)	No. electronic charges/cm² $\times 10^{-11}$
4	0.175	1.76	10	0.190	2.56
5	0.170	1.79	12	0.183	2.60
7	0.162	1.90	14	0.175	2.63
10	0.146	2.07	15	0.173	2.65
12	0.134	2.21	16	0.167	2.69
13	0.127	2.29	18	0.156	2.76
14	0.115	2.36	20	0.143	2.85
15	0.102	2.42	25	0	3.10
16	0.095	2.49			
17	0	2.56			

the valence band. In curves A and B, the number of these states was $N=1.18\times10^{11}$ cm^{-2} and 6.90×10^{10} cm^{-2} respectively. The fit was as good as could be expected. Small ϕ_s values had a relatively larger experimental error because the conductances were very low and therefore not accurate. The number of states at the germanium-germanium oxide interface were so low that they could not have been of the Tamm[3] type, in which the number of states correspond exactly to the number of surface atoms (or multiples thereof). The states which were observed must have arisen from impurity atoms or crystalline defects at the surface.

The variation of the time constant can be understood qualitatively by assuming that the transfer of electrons from the semiconductor to the traps was hindered by the large potential wall in the oxide film. The electrons had to move into the conduction band of the oxide film. The electric field in the oxide film made the potential wall even higher (see Fig. 1, reference 1). It is obvious that the greater the negative charge outside the oxide and the thicker the film, the higher the barrier. Under steady-state conditions, the charge increased with increasing voltage, and the samples which had been exposed for a longer time to wet O_2 and O_3 had thicker oxide films. This model is in qualitative agreement with the observed time constants.

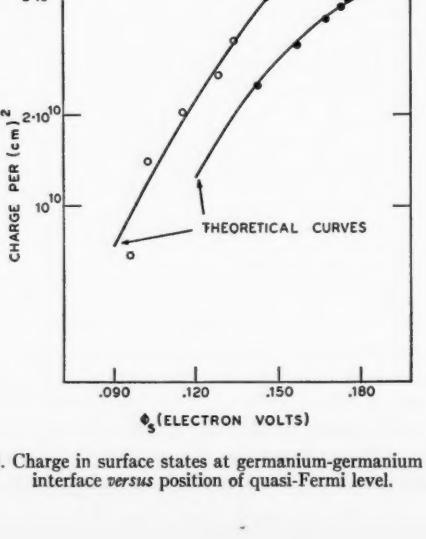


FIG. 3. Charge in surface states at germanium-germanium oxide interface *versus* position of quasi-Fermi level.

* This work has been supported in part by the U. S. Signal Corps

[1] deMars, Statz, and Davis, preceding Letter [Phys. Rev. 98, 539 (1955)].

[2] W. H. Brattain and J. Bardeen, Bell System Tech. J. 32, 1. (1953).

[3] I. Tamm, Physik. Z. Sowjetunion 1, 733 (1932).

Spin-Lattice Relaxation Time of the Si^{29} Nucleus in Pure Fused Silica

G. R. HOLZMAN, JOHN H. ANDERSON, AND W. KOTH
Mellon Institute, Pittsburgh, Pennsylvania

(Received February 21, 1955)

THE Si^{29} nuclei in a sample of Corning purified fused silica (ultraviolet quality) were observed to approach equilibrium magnetization exponentially in a field of 9.2 kilogauss at room temperature, with a relaxation time of 10 hours. This time is a mean of two measurements which differ by approximately 1 percent. Measurement of the relaxation time in the earth's field at room temperature yielded a value of 1.3 min. An emission spectral analysis at Mellon Institute on a similar Corning sample gave 120–530 parts per million for the summed paramagnetic impurities.

The measurement procedure used was to record the change in magnitude of the component of the macroscopic magnetic moment in the plane of and coincident with the rotating rf field vector. This change was brought about by immersing the sample in the magnetic field for varying time intervals, and made evident using a Bloch crossed-coil and phase-sensitive detection system (a Varian spectrometer) with an rms radio-frequency magnetic field intensity of 1.3 gauss. The shape of the nuclear magnetic resonance curve was indicative of Bloch's limiting case of rapid passage through resonance and did not change when the modulation amplitude was varied from approximately 0.05 to 0.5 rms gauss. Points on the exponential curve used in determining this unusually long relaxation time were obtained by measuring the peak-to-peak value of the recorded signal.

When the magnetic field is maintained at the resonance value and the rf field intensity is of the order of one gauss, the signal becomes undetectable after approximately one minute. Signal intensity is lost considerably faster at lower rf fields.

An investigation of the temperature and field dependence of T_1 is now being made, together with a study of the effect of impurities and induced color centers on this relaxation time. Audio-frequency spectroscopy is also contemplated.

Resistivity of Dilute Alloys*

D. L. DEXTER
Institute of Optics, University of Rochester, Rochester, New York
(Received February 15, 1955)

DILUTE substitutional alloys, such as Zn in Cu, have a temperature-independent component to their electrical resistivity which is proportional to the solute concentration. The attempt is usually made[1] to describe the change in resistivity per atomic percent addition, ρ_0, with the following parameters,

$$\rho_0 = a(\Delta Z)^2 + b, \tag{1}$$

where ΔZ is the valence difference between the host material and the substitutional impurity, and where a and b are quantities determined by properties of the two atoms not involving their valence difference; and are generally treated as being independent of position in a given row of the periodic table. It is sometimes considered that this form is made plausible by the existing simple treatments of the dilute alloy problem,[2] in which are treated separately the two cases where (1) a valence difference exists, the scattering potential is taken to be a shielded Coulomb potential, and only the first term in Eq. (1) appears, and (2) the valence difference is zero, and only the second term in Eq. (1) occurs. In cases (1) and (2), the parameters a and b, respectively, must of course be positive quantities. Then the argument is made that in general both effects will occur simultaneously, so that one should add the two terms. It seems to have been overlooked that even in the simplest case interference occurs between these two types of scattering, that a linear term in ΔZ will occur in general, and that in fact it may be expected to be sizable unless one of the other terms is also small.

That this is the case may be seen at once from the procedure of calculating a resistivity change, which is proportional to an average of the square of the matrix element of the scattering potential between the initial and final states:

$$\rho_0 = C\langle (i|V|f)^2 \rangle_{Av}. \tag{2}$$

(The averaging process involves some angular weighting factors which need not concern us here.) If the scattering potential V contains a term V_1 proportional to the valence difference, such as $(e\Delta Z/r)\exp(-qr)$, and also another term V_2 independent of ΔZ, determined by inner potentials and lattice strains, the matrix element of V will be equal to the sum of the matrix elements of V_1 and V_2. Thus the square of the matrix element of V is not given by the sum of the squares of the two parts, as is implied by Eq. (1), but contains an interference term proportional to ΔZ which may distinguish between atoms of higher and lower valence.

$$(V)^2 = (V_1)^2 + (V_2)^2 + 2(V_1)(V_2). \tag{3}$$

This term [the last in Eq. (3)] has a magnitude at least as great as the smaller of the other two, and may have either sign, so that the interference term may add to or subtract from the resistivity.

Averaging Eq. (3) over angles, we obtain $\langle (V_1)^2 \rangle_{Av}$ and $\langle (V_2)^2 \rangle_{Av}$, which correspond to the first two terms in Eq. (1), plus the linear term $2\langle (V_1)(V_2) \rangle_{Av}$ which may be of either sign. [Since (V_1) and (V_2) will in general have different angular dependences it is possible that the averaging process may tend to decrease the magnitude of the term linear in ΔZ by accidental cancellation, but this effect may only be investigated in specific cases by explicit calculations of the matrix elements involved.] The point made here is that Eq. (1) is not justified by theoretical arguments, and also that it would be surprising to find it capable of interpreting experimental results.

The writer is indebted to Dr. E. I. Salkowitz for a stimulating conversation on this subject.

* Research supported in part by funds from the U. S. Air Force under a contract monitored by Headquarters, Air Research and Development Command, Baltimore, Maryland.

[1] J. Linde, Ann. Phys. **15**, 239 (1932).

[2] N. F. Mott and H. Jones, *The Theory of the Properties of Metals and Alloys* (Oxford University Press, London, 1936), Chap. VII. An extension of these calculations has been performed by D. L. Dexter, Phys. Rev. **87**, 768 (1952). Note, however, that J. Friedel [Phil. Mag. **43**, 153 (1952)] has shown that additional complications sometimes arise so that these simple theories are not always applicable.

Resistivity of Dilute Magnesium Alloys

EDWARD I. SALKOVITZ AND A. I. SCHINDLER
Naval Research Laboratory, Washington, D. C.
(Received February 15, 1955)

LINDE'S rule[1] states that for dilute substitutional alloys the increase in specific resistivity of a pure metal due to small alloying additions is equal to a term proportional to the square of the difference in valence of the two elements concerned, plus a second term, or

$$\Delta\rho/A = a\Delta Z^2 + b, \tag{1}$$

where $\Delta\rho/A$ is the increase in specific resistivity per atomic percent addition, ΔZ is the difference in valence between the solute and solvent atoms, and a and b are constants. For convenience we define $\gamma = \Delta\rho/A$. In Linde's work, the monovalent face-centered-cubic metals, copper, silver, and gold, were the solvents. In these cases, Linde found that the parameters a and b

TABLE I. Portion of periodic table showing incremental increase in resistivity of magnesium due to alloying elements.

Row ↓	I	II	III	IV
2	Li 0.75×10^{-6} Ωcm/at. %			
3			Al 2.0×10^{-6} Ωcm/at. %	
5	Ag 0.75×10^{-6} Ωcm/at. %	Cd 0.68×10^{-6} Ωcm/at. %	In 2.0×10^{-6} Ωcm/at. %	Sn 5.2×10^{-6} Ωcm/at. %
6			Tl 3.2×10^{-6} Ωcm/at. %	Pb 5.2×10^{-6} Ωcm/at. %

for a given solvent did not change provided the additions were made from a given period of the periodic table. A change in the period, however, changed the values of these parameters.

To determine whether the rule is applicable to dilute alloys of divalent hexagonal metals,[2] resistivity measurements were made of a series of homogeneous binary alloys of magnesium containing lithium, silver, cadmium, aluminum, indium, thallium, tin, or lead. The details of the experiment will be discussed in a forthcoming paper.

The pertinent data have been plotted in Fig. 1 with the increase in the specific resistivity of magnesium due to alloying as the ordinate and the concentration of alloying element as the abscissa. All comparisons were made with a sample of 99.97 percent pure magnesium. The results may be summarized as follows:

(a) In general, the specific resistivity increased linearly with alloying content. (It should be noted that while a straight line fits the points for smallest additions of tin, the values obtained for the two highest concentrations depart from this line. This may be due to the approach of the solubility limit of tin in magnesium.)

(b) Solutes from Column I of the periodic table give a line with the smallest value of γ; solutes from Column IV give the largest value.

(c) The value of γ for cadmium additions is the same as that for silver and lithium additions. The average values for γ are listed in Table I.

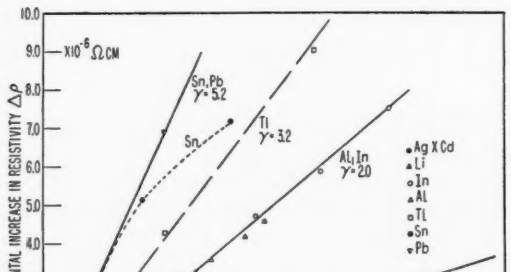


FIG. 1. Incremental increase in resistivity ($\Delta\rho$) of magnesium as a function of composition.

The effect of making additions from elements of the fifth period (silver, cadmium, indium, and tin) may be considered in terms of Eq. (1), recalling that a and b are required to be constants for a given period. Since γ_{Cd} is found to approximately equal γ_{Ag}, then for Linde's rule to be valid, a should equal zero. Physically this implies that a shielded Coulomb field has no effect upon the motion of the conduction electrons, which of course is incorrect. Similarly, since the rule does not consider the sign of ΔZ,[3] Eq. (1) requires γ_{Ag} and γ_{Sn} to be equal. Furthermore, Eq. (1) does not predict the empirical equality of γ_{Li} and γ_{Ag}, of γ_{Al} and γ_{In}, and of γ_{Sn} and γ_{Pb}.

From the above analysis, Linde's rule [Eq. (1)] does not apply to dilute alloys of magnesium. Nor is it apparent that any simple relation may be established analogous to that found by Linde for alloys of the monovalent metals. The data for the alloys of magnesium show, however, that γ depends more significantly upon the column of the periodic table to which the alloying element belongs than upon the period (or row).

We wish to acknowledge helpful discussion with Dr. D. L. Dexter.

[1] J. O. Linde, Ann. Phys. **15**, 239 (1932).
[2] E. I. Salkovitz and A. I. Schindler, Phys. Rev. **91**, 234 (1953).
[3] D. L. Dexter, preceding Letter [Phys. Rev. **98**, 543 (1955)].

Existence of Dirac Generating Functions*

PETER G. BERGMANN
Syracuse University, Syracuse, New York
(Received February 11, 1955)

THIS Letter is the outgrowth of a discussion with V. Bargmann. In a recent paper,[1] it was shown that the generators permissible in a (generalized) Dirac bracket must satisfy the condition

$$\Gamma_{|m}\, U^m{}_{(s)}=0. \tag{1}$$

The question arises whether this requirement is self-consistent. More specifically, given a phase space of $2n$ dimensions and n_0 constraints C^a, of which $2n^*$ are second class, will there exist exactly $2(n+n^*-n_0)$ algebraically independent generators? The answer, to

be proven below, is that this requirement will be satisfied if and only if the first-class constraints form a system of involutions.

The integrability conditions for Eq. (1) on the hypersurface $C^a=0$ are that a small loop whose tangent vector is everywhere a linear combination of null vectors $U^m{}_{(r)}$ can be closed,

$$U^m{}_{(r)|n}\,U^n{}_{(s)}-U^m{}_{(s)|n}\,U^n{}_{(r)}=c_{rs}{}^t\,U^m{}_{(t)}. \quad (2)$$

In the previous paper,[1] it was shown that

$$\begin{aligned} &U^m{}_{(r)}=(C_r,x^m), \quad C_r=U_{(r)a}\,C^a, \\ &U_{(r)a}=\epsilon_{ma}\,U^m{}_{(r)}, \quad U_{(r)a}(C^a,C^b)=0. \end{aligned} \quad (3)$$

Clearly C_r are the first-class constraints, (n_0-2n^*) in number. If we substitute the expression (3) in (2), we obtain:

$$\begin{aligned} (C_r,x^m),_\nu\,\xi^\nu{}_{|n}\,(C_s,x^n)-(C_s,x^m),_\nu\,\xi^\nu{}_{|n}\,(C_r,x^n) \\ =c_{rs}{}^t\,(C_t,x^m) \end{aligned} \quad (4)$$

as our condition of consistency. Working on the left-hand side, we have further:

$$\begin{aligned} &U^m{}_{(r)|n}\,U^n{}_{(s)}-U^m{}_{(s)|n}\,U^n{}_{(r)} \\ &\quad=(C_r,x^m),_\nu[(C_s,\xi^\nu)-(C_s,C^a)\xi^\nu{}_{|a}] \\ &\quad\quad-(C_s,x^m),_\nu[(C_r,\xi^\nu)-(C_r,C^a)\xi^\nu{}_{|a}] \\ &\quad=(C_s,(C_r,x^m))-(C_r,(C_s,x^m))=(x^m,(C_r,C_s)). \end{aligned} \quad (5)$$

The terms $(C_s,C^a)\xi^\nu{}_{|a}$ and $(C_r,C^a)\xi^\nu{}_{|a}$ inside the square brackets of Eq. (5) vanish because the C_s are first-class constraints. Condition (4) will be satisfied if and only if the Poisson bracket between any two first-class constraints is a linear combination of first-class constraints (i.e., free of second-class constraints); in that case, the right-hand side of Eq. (5) will equal a linear combination of null vectors $U^m{}_{(t)}$.

* This work was supported by the Office of Naval Research.

[1] P. G. Bergmann, and I. Goldberg, Phys. Rev. 98, 531 (1955).

Localized Photoeffect in PbS Photocells

E. S. Rittner and S. Fine

Philips Laboratories, Irvington-on-Hudson, New York

(Received February 9, 1955)

SEVERAL years ago an ingenious experiment was proposed by Müser[1] to determine the degree of localization of the photoeffect in semiconductor photocells. It consisted simply in illuminating the photocell with a narrow slit image of high intensity, first parallel and then perpendicular to and overlapping the electrodes. If the photoeffect were confined exclusively to the illuminated region, then for a sufficiently high intensity of illumination the decrease in resistance of the cell would be limited for the parallel orientation by the series resistance of the unilluminated portions, whereas for the antiparallel orientation a much more extensive decrease would occur. On carrying out this experiment on a PbS layer of his own construction (interelectrode distance 1 cm), Müser found substantially the same photoresponse for both orientations of the slit image (width $\frac{1}{3}$ mm), thus implying that the photocarriers produced by the light drifted surprisingly large distances in the field direction.

TABLE I. Photocell resistance *vs* slit orientation and applied field.

Cell No.	R_d, megohms	R_l, megohms Illumination $\perp$	R_l, megohms Illumination $\parallel$ Observed	Computed[a]	E, volts/cm
1	8.29	4.11	6.90	6.90	72.5
	8.31	4.19	6.98	6.93	145
	8.25	4.32	7.03	6.88	290
	8.08	4.36	6.83	6.73	485
2	3.48	1.80	2.36	2.32	145
	3.56	1.83	2.47	2.38	727

[a] The drift of the carriers and the resistance of the illuminated region are neglected. In principle, the actual magnitude of this resistance can be computed from the value of $R_{l\perp}$, if the response is uniform over the photocell. This is hardly the case; however, for the purposes of a rough estimate this procedure indicates the resistance of the illuminated strip to be about 1/12 of the corresponding dark value for cell No.1 and about 1/5 for cell No. 2.

We have repeated this experiment on two PbS photocells obtained from U. S. manufacturers and have obtained results differing from those of Müser. It is believed that this difference has its origin in a lower illumination level and/or a lower sensitivity of the cell studied by Müser.[2] It is a consequence of Ohm's law that for small resistance changes, the response for the two orientations of the light image should be the same. The cogent details of our experiment and the results obtained are presented below.

The photocell was illuminated with a slit image 1 mm in width and of length equal to that of the electrodes. In order to vary the applied field over a wide range without introducing heating difficulties, the cell was placed in series with a pulsed voltage supply of low duty cycle. The currents were measured by varying the resistance of a decade box, also in series with the cell, until the voltage drop across the box, determined oscillographically, attained a predetermined value. In this manner, the dark current and photocurrent were measured with the same field applied to the cell.

Of the two photocells studied, cell No. 1 had an electrode separation of 6 mm and an electrode length of 11 mm; cell No. 2 had an electrode separation of 3 mm and an electrode length of 4.5 mm. Hence, if the photoeffect were localized, we would expect a maximum lowering of the resistance for the parallel orientation of the slit image to a value $\frac{5}{6}$ that of the dark resistance for cell No. 1 and to $\frac{2}{3}$ of the dark value for cell No. 2. The data obtained are presented in Table I. Note that the resistance for the perpendicular orientation of the slit image is appreciably lower than for the parallel orientation and that the observed resistance for the latter case closely approaches the computed limiting value, independent of the magnitude of the applied field. This

implies a highly localized photoeffect, i.e., a drift length for the photocarriers in the field direction small compared to the image width (0.1 cm).

In the case of cell No. 2, the time constant for the decay of the photoresponse is stated by the manufacturer to be 2×10^{-4} second. If it is assumed that this time constant represents the lifetime τ of the free photocarriers, then the expected drift length w for the higher-field case (taking the mobility μ to be of the order of 1 cm²/volt sec)[3] would be

$$w=\mu E\tau=0.15 \text{ cm},$$

a distance large compared to the drift length implied by Table I. The most likely explanation of this discrepancy is that the time constant is very much longer than the free lifetime owing to the phenomenon of trapping[4] in these layers. The importance of trapping has heretofore not received recognition in theoretical treatments of this subject (summarized in reference 3) and complicates greatly the task of deriving an adequate theory.

We are indebted to H. Müser and to J. R. Haynes for helpful exchanges of correspondence relative to this experiment.

[1] H. Müser, Z. physik. Chem. **198**, 52 (1951).
[2] In private correspondence Dr. Müser has expressed his agreement with this explanation.
[3] R. A. Smith, Advances in Physics **2**, 321 (1953).
[4] A. Rose, RCA Rev. **12**, 362 (1951); J. R. Haynes and J. A. Hornbeck, Phys. Rev. **90**, 152 (1953); H. Y. Fan, Phys. Rev. **92**, 1424 (1953); **93**, 1434 (1954).

New Class of Ferroelectrics

A. N. Holden, B. T. Matthias, W. J. Merz, and J. P. Remeika
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received February 24, 1955)

We find that guanidine aluminum sulfate hexahydrate,[1] $(CN_3H_6)Al(SO_4)_2\cdot6H_2O$, is ferroelectric from as high a temperature as it will stand for a short time (about 200°C) down to the temperature of liquid N_2, the lowest we have employed. The material is trigonal, with perfect cleavage in the basal plane; the direction of spontaneous polarization is along the trigonal axis. Ferroelectricity is not usually observed when electrodes are applied to naturally grown faces; the effect can be seen when new faces are generated by abrasion or better by cleavage.

At room temperature the saturation polarization is about 0.35 microcoulombs per sq cm, and the coercive force is between 1200 and 1500 volts per cm measured at 60 cps. These quantities increase with decreasing temperature. The small-signal dielectric constant is about 15 along the axis and about 5 perpendicular to the axis. Of isomorphous substances, we have so far examined those in which D_2O replaces H_2O, in which Ga^{3+} and Cr^{3+} replace Al^{3+}, and in which $(SeO_4)^{2-}$ replaces $(SO_4)^{2-}$. These substances show essentially the same properties. Dr. S. Geller and Dr. E. A. Wood have in progress a determination of the crystal structure of the material; they inform us that the space group is $C_{3v}(2)-P31m$, and the unit cell contains three molecules. The material bears no obvious structural relation to previously known ferroelectrics.

[1] F. Ferraboschi, Proc. Cambridge Phil. Soc. **14**, V, 473 (1908).

Radiation-Controlled Electroluminescence and Light Amplification in Phosphor Films

D. A. Cusano
General Electric Research Laboratory, Schenectady, New York
(Received February 24, 1955)

This is a preliminary report of a new phenomenon observed when voltage is applied to a layer of ZnS:Mn,Cl phosphor during excitation by uv or x-rays. The luminescent layer (about 10 microns thick) is deposited by vapor reaction.[1,2] As contrasted with the granular phosphor commonly used for electrolumines-

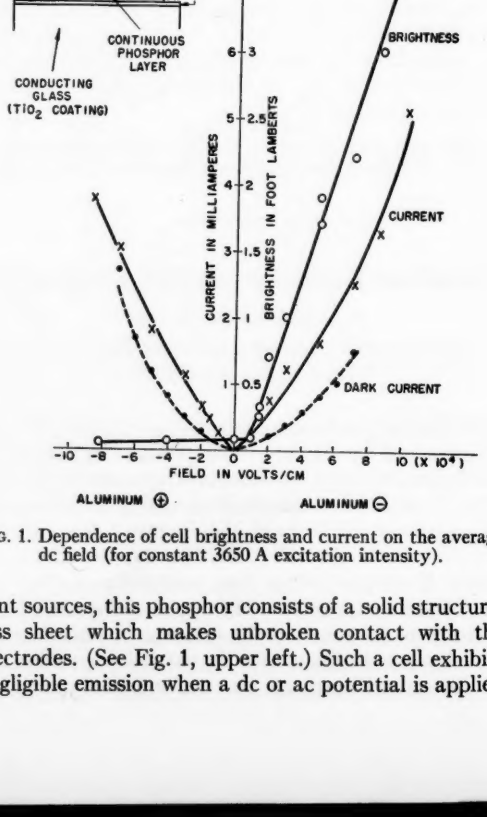


Fig. 1. Dependence of cell brightness and current on the average dc field (for constant 3650 A excitation intensity).

cent sources, this phosphor consists of a solid structureless sheet which makes unbroken contact with the electrodes. (See Fig. 1, upper left.) Such a cell exhibits negligible emission when a dc or ac potential is applied

to the electrodes. However, when the layer is excited by uv or x-rays and emits the yellow luminescence typical of manganese, the simultaneous effect of the electric field increases the brightness markedly. For dc, this occurs when the metal electrode of the cell is negative. The behavior of a typical cell under constant 3650 A excitation is shown as a function of the average field in Fig. 1. Increase in brightness of eightyfold has been observed. Equally large changes have been seen with x-ray excitation. Previous work on granular phosphors involving x-rays and ac fields has been reported by Destriau.[3]

For 3650 A, the dependence of enhancement on excitation intensity is shown in Fig. 2, where the average field is fixed at 10^5 v/cm. In the given range,

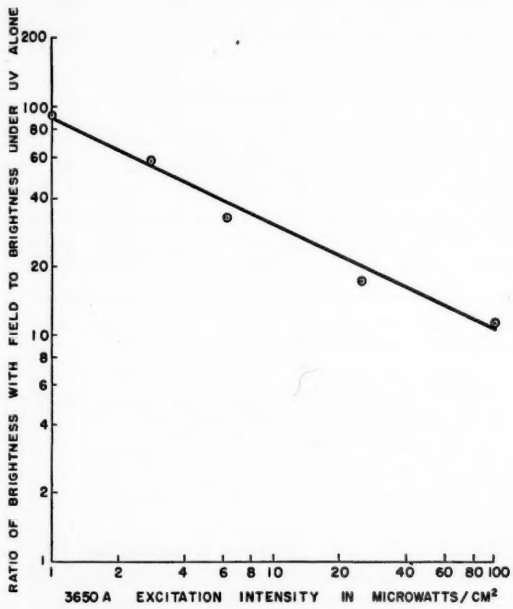


FIG. 2. Dependence of enhancement (i.e., ratio of the brightness with field and uv to that with uv alone) on 3650 A excitation intensity (for fixed dc field = 10^5 v/cm).

the enhancement varies inversely as the square root of the excitation intensity. By computing the visible output in relation to the incident ultraviolet, one finds (for low excitation intensity) as many as 10 photons of visible light are produced for each photon of 3650 A. For higher uv intensities, the photon amplification ratio decreases until it reaches unity in the neighborhood of 100 microwatts/cm². In spite of the wavelength conversion, this represents true light amplification, in the sense that more radiant energy is emitted than is incident on the layer.

The over-all efficiency of the light production is not very high if one considers the energy derived from the voltage source. Neglecting only the dark current, which may be leakage at weak points in the layer, the luminous efficiency is found to be approximately 0.2 lumen/watt. With regard to time constants, the light emission responds n milliseconds to changes in the dc field, provided these changes occur above the threshold shown in Fig. 1. In contrast, the time constants associated with changes in exciting radiation increase with decreasing intensity, approaching two or three seconds in the 1-microwatt/cm² region. The emission spectrum itself is seen to shift approximately 100 A toward shorter wavelength with high field (10^5 v/cm).

More detailed information will be presented in a subsequent paper. The general phenomenon is interpreted as the initiation and control of electroluminescence through the absorption of electromagnetic energy. Based on the theoretical work of Piper and Williams[4] on electroluminescence, Williams[5] earlier suggested the possibility of light-amplifying phosphors.

[1] D. A. Cusano and F. J. Studer, U. S. Patent 2 685 530.
[2] F. J. Studer and D. A. Cusano, J. Opt. Soc. Am. (to be published).
[3] G. Destriau, Electrochemical Society Meeting in Chicago, 1954 (unpublished). M. Destriau, Compt. rend. 238, 2298 (1954).
[4] W. W. Piper and F. E. Williams, Phys. Rev. 87, 151 (1952); Brit. J. Appl. Phys., Supplement 4, 39 (1955).
[5] F. E. Williams, following Letter [Phys. Rev. 98, 547 (1955)].

Theoretical Basis for Light-Amplifying Phosphors

FERD E. WILLIAMS
General Electric Research Laboratory, Schenectady, New York
(Received February 24, 1955)

THE electroluminescence of ZnS:Cu single crystals has been interpreted on the basis of a mechanism involving the creation of charge carriers in a high-field region where acceleration of the charge carriers to kinetic energies sufficient for the excitation or ionization of activator systems by inelastic collisions occurs. The charge carriers are created by field ionization of deep donor levels in an exhaustion layer formed at the cathode contact by ionization of shallow donors.[1] The series of sequential events required for electroluminescence suggested several years ago the possibility of control of electroluminescence by incident radiation and of radiation amplification in phosphors. Specifically the creation of charge carriers in a high-field region by photon absorption, rather than by field ionization of impurity systems, followed by acceleration and multiplication in the high-field region was recognized as a feasible mechanism for light amplification. Amplification depends on sufficient local field so that either the charge carriers created by the incident radiation attain sufficient kinetic energy for the excitation by impact of the activator systems several times in their mean lives,

or the charge carriers created by the incident radiation form avalanches of prebreakdown magnitudes and excitation by impact of a number of activators per incident photon is attained. The charge carrier multiplication is attainable with the high fields characteristic of electronic breakdown. Rupture may be avoided by the high-field region being sufficiently thin so that the required number of multiplications for breakdown cannot occur.

The alternative electroluminescent mechanism of minority charge carrier injection followed by capture by appropriate activator systems also conceptually provides a method of radiation amplification in a phosphor.[2] In this case, the photons create charge carriers which are trapped, thereby yielding a space charge facilitating the transport of carriers of opposite sign which are subsequently captured by the activator systems resulting in ionization. Capture of the majority charge carrier then leads to photon emission, conceivably in greater quantity than the initial photon absorption.

The observations of Cusano[3] on thin films of ZnS:Mn,Cl are in accord with the impact excitation mechanism of radiation amplification. The energy levels of the Mn activator system are deep-lying, below the valence band of ZnS, and therefore, collision excitation rather than the capture of injected charge carriers must be involved. ZnS:Mn,Cl, in contrast with electroluminescent ZnS:Cu, is relatively free of deep donors. The Cl coactivator introduces electron traps approximately 0.3 ev below the conduction band. The absorption of 3650 A radiation apparently excites valence electrons to these states and the coactivator can thus behave as a donor in a high-field region. The rectification of the light-amplifying phosphor film is not unexpected. The Piper-Williams theory[1] emphasized that a cathode barrier, probably of the exhaustion type, exists at the metal-semiconductor contact for electroluminescent ZnS:Cu single crystals. A similar barrier probably exists at the metal-semiconductor contact of the ZnS:Mn,Cl films; and electroluminescence, initiated by 3650 A radiation, occurs only when the metal electrode is negative. The conditions at the ZnS—TiO_2 interface are less apparent. Although the average field through the film during light amplification is 10^5 v/cm, the field in the cathode barrier is probably 10^6 v/cm. In fact, the 100 A shift in the emission spectrum observed under these conditions of excitation can be accounted for by a field of this magnitude perturbing the Mn activator system essentially by the Stark effect. Fields of 10^5 to 10^6 v/cm are also indicative of the collision excitation mechanism of electroluminescence.

[1] W. W. Piper and F. E. Williams, Phys. Rev. **87**, 151 (1952); Brit. J. Appl. Phys., Supplement No. 4, 39 (1955).
[2] W. W. Piper and F. E. Williams, Phys. Rev. (to be published).
[3] D. A. Cusano, Bull. Am. Phys. Soc. **30**, No. 1, 30 (1955); and preceding Letter [Phys. Rev. **98**, 546 (1955)].

Anomaly in the Low-Temperature Atomic Heat of Silver*

P. H. KEESOM AND N. PEARLMAN
Department of Physics, Purdue University, Lafayette, Indiana
(Received February 15, 1955)

SINCE the publication of an earlier Letter[1] with this title, Clement,[2] analyzing the effect of deviations of earlier temperature scales from the 48 scale, pointed out that deviations of the adopted scale from the true thermodynamic scale may cause apparently significant anomalies in calorimetric results. The recent work of Erickson and Roberts[3] and of Keller[3] substantiate the deductions of Kistemaker[3] concerning deviations from the 48 scale, and a new scale, for convenience called the 55 scale, seems to be established. We recalculated the specific heat data based on the adopted 48 scale, using the formula

$$c_{55}=c_{48}[1+d(T_{48}-T_{55})/dT].$$

In plotting c_{48}/T *versus* T^2, it was not possible to draw one straight line through the data and an anomaly appeared significant. However, this anomaly disappears if the 55 scale is used and one straight line is able to represent the data within their accuracy. Using least squares, we now obtain for the atomic heat of silver below 4°K: $c=0.66T+0.170T^3$ millijoules/mole deg, which corresponds to a Debye parameter θ_0 of 225°K.

* Assisted by Signal Corps.
[1] P. H. Keesom and N. Pearlman, Phys. Rev. **88**, 140 (1952).
[2] J. R. Clement, Phys. Rev. **93**, 1420 (1954).
[3] See W. E. Keller, Phys. Rev. **97**, 1 (1955).

Propagation and Plasma Oscillation in Semiconductors with Magnetic Fields*

BENJAMIN LAX AND LAURA M. ROTH†
Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts
(Received February 28, 1955)

MAGNETO-IONIC theory[1] can be extended to semiconductors whose energy surfaces in the Brillouin zone are ellipsoidal[2] or warped spheres[3] possessing cubic symmetry. With a magnetic field **B** along the [001], [110], or [111] axes, the rf conductivity tensors for electrons in germanium and silicon can be written

$$\sigma=\begin{bmatrix}\sigma_{11} & \sigma_{12} & 0\\ -\sigma_{12} & \sigma_{22} & 0\\ 0 & 0 & \sigma_{33}\end{bmatrix}, \qquad (1)$$

where the z axis is along **B**. For the [001] and [111] directions $\sigma_{11}=\sigma_{22}$. The conductivity tensors, obtained for each of the eight or six ellipsoids in their respective

coordinate systems, are transformed to the coordinate system of Eq. (1) and added to yield the above result.

Maxwell's equations yield a solution for the propagation constant Γ given by

$$\Gamma^2 = -\omega^2\epsilon\mu_0 + j\omega\mu_0\sigma_{eff}, \quad (2)$$

where ω is the frequency of an rf field $\mathbf{E}$, ϵ is the dielectric constant, and σ_{eff} becomes

$$\alpha\|\mathbf{B}: \quad \sigma_{eff} = \tfrac{1}{2}(\sigma_{11}+\sigma_{22}) \pm j[\sigma_{12}^2 - \tfrac{1}{4}(\sigma_{11}-\sigma_{22})^2]^{\frac{1}{2}}, \quad (3a)$$

$$\alpha\perp\mathbf{B}: \begin{cases} \sigma_{eff} = \sigma_{33}, & (3b) \\ \sigma_{eff} = \dfrac{\omega\epsilon(\alpha_2^2\sigma_{11}+\alpha_1^2\sigma_{22}) - j(\sigma_{12}^2+\sigma_{11}\sigma_{22})}{\omega\epsilon - j(\alpha_1^2\sigma_{11}+\alpha_2^2\sigma_{22})}, & (3c) \end{cases}$$

where $\boldsymbol{\alpha}$ is the unit vector along $\boldsymbol{\Gamma}$. For longitudinal propagation, Eq. (3a), $\mathbf{E}$ is elliptically polarized normal to $\mathbf{B}$ and circularly polarized when $\sigma_{11}=\sigma_{22}$. Transverse propagation has two cases, (3a) corresponding to linear polarization with $\mathbf{E}$ along $\mathbf{B}$ and (3b) to elliptical polarization with $\mathbf{E}\perp\mathbf{B}$. The effective conductivities for $\mathbf{B}$ along the [001] direction for the eight and the [111] for the six ellipsoids are identical, and for $\boldsymbol{\alpha}\|\mathbf{B}$ become

$$\sigma_{eff}/\sigma^* = [1 \pm (p+2)jb/(2p+1)]/[1+b^2(p+1)/3p]. \quad (4)$$

Linear polarization for $\boldsymbol{\alpha}\perp\mathbf{B}$ yields

$$\sigma_{eff}/\sigma^* = [1+3b^2/(2p+1)]/[1+(p+2)b^2/3p], \quad (5)$$

where $\sigma^* = ne^2/m^*(\nu+j\omega)$, $m^* = 3m_1m_2/(2m_1+m_2)$, $p=m_1/m_2$, and $b=eB/m_2(\nu+j\omega)$, m_1, m_2 being the longitudinal and transverse components of the mass tensor, and ν the collision frequency, assumed independent of energy. The expressions for σ_{eff} with $\mathbf{B}$ along other principal directions are more involved except for the linear polarization, where for the eight ellipsoids, $\mathbf{B}$ along [111] and [110] directions, the conductivities become, respectively,

$$\frac{\sigma_{eff}}{\sigma^*} = \frac{[1+(7p+2)b^2/3p(2p+1)]}{[1+(p+8)b^2/9p]}; \quad (6a)$$

$$\frac{\sigma_{eff}}{\sigma^*} = \frac{[1+(p+2)b^2/(2p+1)]}{[1+(2p+1)b^2/3p]}. \quad (6b)$$

For the six ellipsoids with $\mathbf{B}$ along [110] and [001] axes, the conductivities become, respectively,

$$\frac{\sigma_{eff}}{\sigma^*} = \frac{[1+(p+5)b^2/2(2p+1)]}{[1+(p+1)b^2/2p]}; \quad (7a)$$

$$\frac{\sigma_{eff}}{\sigma^*} = 1. \quad (7b)$$

Setting $\Gamma=0$ gives equations for plasma oscillation frequencies. These become real if ν is neglected, i.e., for "pure" samples, low temperatures, and high frequencies. $\mathbf{E}\|\mathbf{B}$ results in linear and $\mathbf{E}\perp\mathbf{B}$ in elliptical (or circular if $\sigma_{11}=\sigma_{22}$) polarization. The equations, identical for $\mathbf{B}$ along the [111] and [100] directions for the six and eight ellipsoid models respectively, are

$$\omega^4 - \omega^2[(p+2)\omega_2^2/3p + \omega_p^2] + 3\omega_p^2\omega_2^2/(2p+1) = 0, \quad (8a)$$

$$\omega^3 - \omega[(p+2)\omega_2^2/3p + \omega_p^2] \pm (p+2)\omega_p^2\omega_2/(2p+1) = 0, \quad (8b)$$

where $\omega_p^2 = ne^2/m^*\epsilon$ and $\omega_2 = eB/m_2$.

Equations (8a) and (8b) describe the linear and circular polarizations respectively. With $\mathbf{E}\|\mathbf{B}$ and along the [100] axes, for six ellipsoids, there is no magnetic effect and $\omega^2=\omega_p^2$. Other cases will be discussed in a later publication.

The above results depend on crystal orientation and parameters related to energy surface curvature and may be utilized for investigating the shape of energy bands in anisotropic semiconductors at microwave frequencies. The approach outlined may also be applicable for analyzing the anisotropy of warped spheres.[3] Approximate expressions of the conductivity tensor for the latter are being derived from Boltzmann transport theory by Zeiger.[4] Actual experiments require finite samples and depolarizing effects may have to be taken into account as suggested by Kittel.[5] The conductivities would then contain components of a depolarizing tensor.

* The research in this document was supported jointly by the Army, Navy, and Air Force under contract with the Massachusetts Institute of Technology.

† Now at Harvard University, Cambridge, Massachusetts.

[1] E. V. Appleton, U.R.S.I. Reports, Washington (1927).

[2] S. Meiboom and B. Abeles, Phys. Rev. 93, 1121 (1954); Lax, Zeiger, Dexter, and Rosenblum, Phys. Rev. 93, 1418 (1954); Dexter, Lax, Kip, and Dresselhaus, Phys. Rev. 96, 222 (1954).

[3] Dexter, Zeiger, and Lax, Phys. Rev. 95, 557 (1954); R. N. Dexter and B. Lax, Phys. Rev. 96, 223 (1954).

[4] H. J. Zeiger (private communication).

[5] C. Kittel, Physica (to be published).

Space Charge Layer Near the Surface of a Ferroelectric

WERNER KÄNZIG

Department of Physics, University of Illinois, Urbana, Illinois

(Received February 25, 1955)

IT is known that a space charge layer can exist near the surface of solids. In the case of a *semiconductor*, this layer arises from the trapping of electrons in the surface traps.[1] The electric potential V between the surface and the bulk is about 1 volt, and the thickness of the layer is generally between 10^{-4} cm and 10^{-6} cm. Hence, electric fields of the order of magnitude 10^4 volts/cm to 10^6 volts/cm may prevail. In the case of an *ionic conductor*, the space charge layer is due to the fact that the surface is a source or a sink of vacancies.[2] V is again of the order of magnitude 1 volt. The thickness of the layer decreases with increasing vacancy concentration. Thin space charge layers ($<10^{-5}$ cm) require a vacancy concentration which exceeds 10^{17}

cm^{-3}. In *intrinsic* ionic conductors this concentration is generally reached only at rather elevated temperatures. In *extrinsic* ionic conductors, however, an impurity concentration near 10^{18} cm^{-3} leads to the formation of what we might call an "ionic Schottky exhaustion layer." Its thickness lies between 10^{-5} and 10^{-6} cm and does not depend sensitively upon the temperature.[3] Hence, fields as high as 10^5 volts/cm to 10^6 volts/cm may result.

An interesting effect should arise from these space charge layers in the case of highly polarizable crystals, such as ferroelectrics above the Curie point. Fields of 10^5 to 10^6 volts/cm induce the saturation polarization. The resulting piezoelectric or electrostrictive strain is about 1 percent and thus would be easily observable. [The fact that the thickness of the layer is proportional to $\sqrt{\epsilon}$, where ϵ is the static dielectric constant, should not prevent this effect: by inserting the dielectric constant of the saturated crystal ($\epsilon \approx 200$) one obtains a layer which is still thin enough to justify the assumption of dielectric saturation.] To summarize, we can say that a ferroelectric semiconductor or a ferroelectric extrinsic ionic conductor can exhibit a surface layer of a thickness between 10^{-5} cm and 10^{-6} cm where saturation polarization and the corresponding piezoelectric strain prevails. This layer does not depolarize if the crystal is heated above the Curie temperature of the bulk because the properties of the space charge layer do not depend sensitivity upon the temperature as long as the conditions for dielectric saturation are fulfilled.

There is considerable experimental evidence that such a layer exists in the case of ferroelectric $BaTiO_3$ which is known to be a semiconductor[4] and which might also be an extrinsic ionic conductor to some extent. Diffusion measurements indicate that intrinsic ionic conduction is negligible in the temperature range which is of interest here.[5] X-ray and electron diffraction experiments with very small $BaTiO_3$ particles have led to the following conclusions[6]:

1. There is a discrepancy in symmetry between a surface layer of a thickness of about 100 A and the bulk. The misfit between the surface layer and the bulk is smaller below the Curie temperature of the bulk. Thus the structure of the surface layer is closer to the tetragonal structure of the polarized lattice than to the cubic structure of the unpolarized lattice.

2. Electron diffraction experiments indicate a tetragonal strain in the surface layer which is slightly larger than the tetragonal strain of the bulk below the Curie temperature. The tetragonal surface strain does not vanish if the crystal is heated above the Curie temperature of the bulk.

It is interesting to note that in the case of ferroelectric KH_2PO_4 no evidence for a polarized surface layer was found.[7] As KH_2PO_4 is not a semiconductor and generally much purer than $BaTiO_3$, this supports the view that the observed surface strain in the case of $BaTiO_3$ is due to an ionic or electronic space charge layer. This layer may also account for the observation that very thin, virgin $BaTiO_3$ crystals do not show normal dielectric hysteresis at low voltages V, though the apparent electric field (calculated as V/d, where d is the thickness of the crystal) is larger than the coercive field. The space charge layer has to be broken down before the normal dielectric hysteresis can be observed.[8]

The author is indebted to J. Bardeen, R. J. Maurer, and F. Seitz, for illuminating discussions.

[1] John Bardeen, Phys. Rev. **71**, 717 (1947).
[2] J. Frenkel, *Kinetic Theory of Liquids* (Clarendon Press, Oxford, 1947).
[3] K. Lehovec, J. Chem. Phys. **21**, 1223 (1953).
[4] Busch, Flury, and Merz, Helv. Phys. Acta **21**, 212 (1948); S. Nomura and S. Sawada, J. Phys. Soc. Japan **5**, 227 (1950); E. K. Weise and J. A. Lesk, J. Chem. Phys. **21**, 801 (1953).
[5] A. Garcia-Verduch and R. Lindner, Ark. Kemi **5**, 313 (1953).
[6] Anliker, Brugger, and Känzig, Helv. Phys. Acta **27**, 99 (1954).
[7] Jaccard, Känzig, and Peter, Helv. Phys. Acta **26**, 521 (1953).
[8] W. J. Merz, Phys. Rev. **76**, 1221 (1949).

Diamagnetic Resonance in Electronic Conductors

R. B. DINGLE

Department of Physics, University of Western Australia, Nedlands, Australia

(Received February 4, 1955)

IN a recent Letter with the above title,[1] Dorfman points out that he had published an article in 1951 predicting the general features of diamagnetic ("cyclotron") resonance in electronic conductors, which would therefore be prior to my detailed theoretical treatment[2] appearing in 1952.

In order to avoid any possible misunderstandings, I should like to draw attention to the fact that I had described the effect at the 1951 International Conference on Very Low Temperatures, and (like Dorfman) had emphasized that it would lead to a direct estimate of the effective mass of the carriers. The publications of such general descriptions by myself[3] and Dorfman[4] were therefore independent and essentially simultaneous. The detailed treatment[2] appearing in the *Proceedings of the Royal Society* (*London*) for 1952 was taken from my 1951 doctorate thesis.[5]

[1] J. G. Dorfman, Phys. Rev. **96**, 1704 (1954).
[2] R. B. Dingle, Proc. Roy. Soc. (London) **A212**, 38 (1952).
[3] R. B. Dingle, Proceedings of the International Conference on Very Low Temperatures, edited by R. Bowers (Oxford, 1951), abstract on p. 165.
[4] J. G. Dorfman, Doklady Akad. Nauk. S.S.S.R. **81**, 765 (1951). (Unfortunately, this volume does not appear to have become available in England until after the publication of my 1952 paper.)
[5] R. B. Dingle, doctorate thesis, Cambridge University (1951), Chap. VI.

Heat Capacity and Entropy of Liquid He³ from 0.23 to 2°K: Nuclear Alignment in Liquid He³†‡

BERNARD M. ABRAHAM, DARRELL W. OSBORNE, AND BERNARD WEINSTOCK
Argonne National Laboratory, Lemont, Illinois

(Received January 26, 1955; revised manuscript received February 10, 1955)

WE have extended our measurements[1] of the heat capacity of liquid He³ to cover the temperature range 0.23 to 2°K, and the results are shown in Fig. 1, along with the values reported by Roberts and Sydoriak.[2] The data of deVries and Daunt[3] have not been included because of their very large scatter. Also shown are calculated curves of C_V for an ideal Fermi-Dirac gas with degeneracy temperatures of 0.45°K (suggested by Fairbank, Ard, and Walters[4] to fit their susceptibility measurements) and 4.98°K (the Fermi temperature for He³ at the liquid density). A comparison of these curves with the experimental measurements[5] shows that the ideal Fermi-Dirac gas model is unsatisfactory for liquid He³.

The entropy of liquid He³ between 0.23 and 2°K was calculated from the equation

$$S=2.52+\int_{1.5}^{T} C_{sat} d\ln T,$$

where the constant 2.52 cal deg^{-1} mole^{-1} is the entropy of the liquid at 1.5°K obtained from our vapor pressure data[6] and from new values of the second virial coefficient[7] and the liquid density.[8] The entropy obtained in this way is given in Fig. 2 (curve A), together with the values at 1, 1.5, and 2°K calculated from the vapor pressure. Since there is an estimated uncertainty of 0.17 cal deg^{-1} mole^{-1} in the entropy at 1.5°K, the whole entropy curve may have to be shifted up or down when an accurate measurement of the heat of vaporization is obtained.

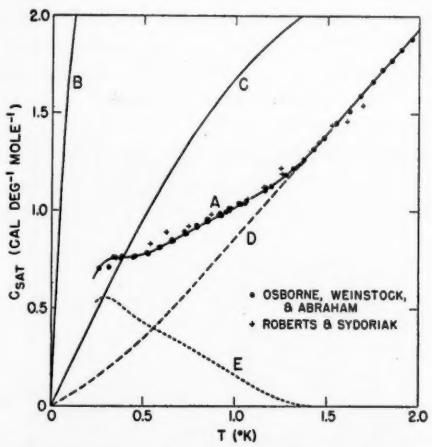


FIG. 1. Heat capacity of liquid He³ at its saturated vapor pressure. (A) Experimental results; (B) and (C), C_V of ideal Fermi-Dirac gas with degeneracy temperatures of 0.45° and 4.98°K, respectively; (D) estimated nonspin heat capacity; (E) estimated spin heat capacity (curve A minus curve D).

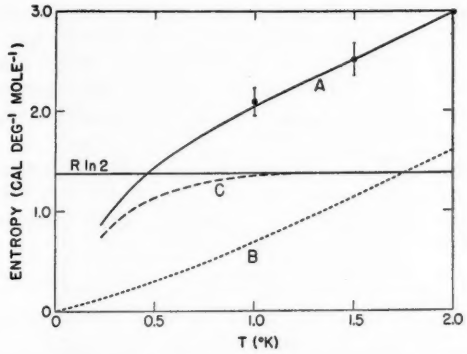


FIG. 2. Entropy of liquid He³. The circles show the entropy calculated from the vapor pressure. Curve A gives the entropy calculated from the heat capacity data and the value 2.52 cal deg^{-1} mole^{-1} for the entropy at 1.5°K. (B) Estimated nonspin entropy [from integration of curve D], Fig. 1; (C) estimated spin entropy, S_{spin}, [curve A minus curve B].

The entropy at 0.23°K is 0.86 cal deg^{-1} mole^{-1}, a value significantly smaller than $R\ln 2=1.38$ cal deg^{-1} mole^{-1}, the entropy for random orientation of the nuclear spins. It follows, without any assumption regarding other contributions to the entropy, that the nuclear spins in liquid He³ must be appreciably ordered at this temperature. A similar conclusion was drawn by Roberts and Sydoriak[2] from their preliminary value of 1.20 cal deg^{-1} mole^{-1} for the entropy at 0.4°K.

Fairbank, Ard, and Walters have demonstrated nuclear alignment in liquid He³ more directly by measurement of the nuclear magnetic susceptibility.[4] In order to make a quantitative comparison between our results and those of Fairbank, Ard, and Walters we have empirically resolved the heat capacity of liquid He³ into two parts, one attributed to the alignment of the nuclear spins and the other to the remaining types of excitation. A striking similarity was noticed between the heat capacity of liquid He⁴ above 2.5°K[9] and that of liquid He³ above 1.4°K, in that both heat capacities vary linearly with temperature and extrapolate nearly to zero. It seems reasonable to interpret this "high temperature" heat capacity as arising from excitations in which statistics and spin do not play an important role. The linear part of the heat capacity curve was extrapolated to 0°K (Fig. 1, curve D) and was assumed to represent the nonspin contribution to the heat capacity. The difference between this curve and the total heat capacity gives the spin heat capacity, which is plotted in Fig. 1 as curve E. The entropy was resolved by integrating the nonspin heat capacity from 0°K to obtain the nonspin contribution to the entropy (Fig. 2, curve B) and subtracting this from the total

entropy to obtain the spin contribution (Fig. 2, curve C). Actually the extrapolation of the nonspin heat capacity was not strictly linear but rather was curved below 0.75°K so that it would pass through the origin and also would integrate to give R ln2 for the spin entropy at 1.4°K.

The fraction of the spins which are not aligned was then calculated from the ratio $S_{spin}/(R \ln 2)$.[10] These values were found to be in satisfactory agreement with the measurements of $\chi T/C$.[4]

† This work was performed under the auspices of the U. S. Atomic Energy Commission.

‡ Presented as an invited paper at the meeting of the American Physical Society at Chicago, Illinois, November 26, 1954.

[1] Osborne, Abraham, and Weinstock, Phys. Rev. **94**, 202 (1954).
[2] T. R. Roberts and S. G. Sydoriak, Phys. Rev. **93**, 1418 (1954).
[3] G. deVries and J. G. Daunt, Phys. Rev. **92**, 1572 (1953); **93**, 631 (1954).
[4] Fairbank, Ard, and Walters, Phys. Rev. **95**, 566 (1954).
[5] Strictly speaking, the difference between C_{sat} and C_V should be taken into account before making this comparison. Although data are not available for a precise calculation of this difference, the correction is in the direction of making the discrepancy even greater than that shown.
[6] Abraham, Osborne, and Weinstock, Phys. Rev. **80**, 366 (1950).
[7] Kilpatrick, Keller, Hammel, and Metropolis, Phys. Rev. **94**, 1103 (1954).
[8] E. C. Kerr, Phys. Rev. **96**, 551 (1954).
[9] W. H. Keesom, *Helium* (Elsevier Publishing Company, Amsterdam, 1942), p. 212.
[10] L. Goldstein, Phys. Rev. **96**, 1455 (1954), has independently suggested a similar resolution of the heat capacity and entropy of liquid He^3 and has presented a theoretical basis for the relation $S_{spin}=(\chi T/C)R \ln 2$.

Effect of Oxygen Gas on the Photoconductivity of BaO

M. Sakamoto, S. Kobayashi, and S. Ishii
Tokyo College of Science, Kagurazaka, Shinjuku-ku, Tokyo, Japan
(Received February 15, 1955)

When a barium-oxide coated cathode with a probe electrode, as shown in Fig. 1, is irradiated by ultraviolet light, a remarkably steep peak is observed at about 3.7 ev.[1,2] The author has proposed that this peak might be due to a level of surface origin, since it was observed only in powders and not in single crystals.[2] To test this idea, we have observed the change of the photocurrents before and after the breaking of an oxygen ampule, which was inserted in the vacuum tube. These observations were made especially at the 3.7-ev peak and at the 4.2-ev hump which was thought to be due to the thermal dissociation of the excitons.[3] According to the results shown in Fig. 2, the photocurrent at 3.7 ev decreases rapidly with a time constant of several tens of seconds after exposure to the oxygen gas and finally reaches about 20 percent of the initial value. On the other hand, the photocurrent at 4.2 ev decreases slowly with a time constant of about 5 minutes, and the final total decrease is only about 50 percent. The author has repeated this experiment with argon gas, but did not find in this case such a difference in behavior between the 3.7-ev and 4.2-ev photocurrents.

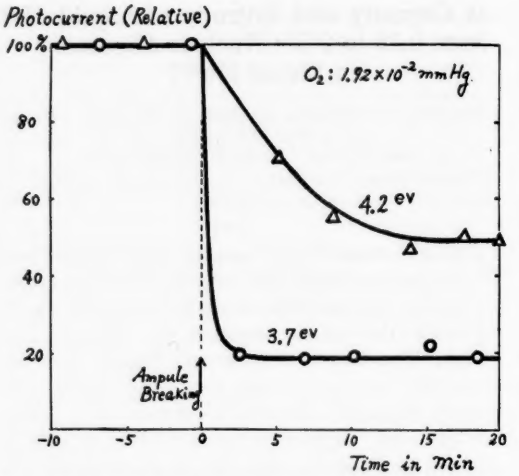


Fig. 2. The effect of oxygen gas on the photoconductivity of BaO. The relative photocurrents are shown to facilitate comparison.

Therefore, the decay phenomena of the two currents are of essentially different types. As one possibility, the decay at 4.2 ev may be interpreted as the scattering of the conduction electrons by the oxygen molecules which diffused into pores existing between the powdered small crystals. The calculated value of the time constant for the simple model shown in Fig. 3 (Knudsen's formula) was reasonable in comparison with the observed value. The decay at 3.7 ev may, on the other hand, be considered as supporting the proposition of a surface level.

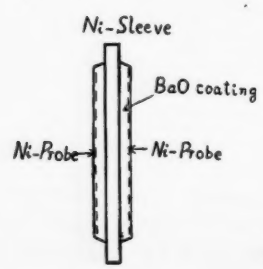


Fig. 1. BaO cathode specimen.

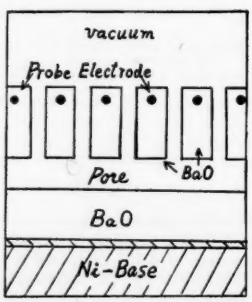


Fig. 3. Simple model for the calculation.

The author wishes to express his thanks to Dr. H. Kawamura, Dr. A. Kobayashi, and M. Tomura for profitable discussions.

[1] H. B. DeVore and J. W. Dewdney, Phys. Rev. **83**, 805 (1951).
[2] M. Sakamoto, Le Vide, **52–53**, 109 (1954).
[3] W. W. Tyler and R. L. Sproull, Phys. Rev. **83**, 548 (1951).

X-Ray Analysis of Compressed KCl Powder Crystals

J. G. KEREIAKES*
University of Cincinnati, Cincinnati, Ohio
(Received December 20, 1954; revised manuscript received February 14, 1955)

IN view of the wide use of compressed crystals in infrared transmission studies and of the possible use of compressed crystals as the crystal phosphor in scintillation counting, it is important that the physical effects of high pressure on powder crystals be investigated. Presented here are x-ray diffraction patterns of compressed potassium chloride powder crystals.

The compressed crystal samples were obtained as follows. About 0.1 gram of 60–80 micron grain-size KCl powder was placed on the surface of an oil-hardened chrome-nickel alloy die. After sealing, the die and powder containing compartment was evacuated until approximately 10^{-3} mm Hg pressure was attained. The powder samples were subjected to pressures ranging from 5000 lb/in.^{2} to 200 000 lb/in.^{2} The resulting samples were in the form of hardened disks of $\frac{1}{2}$ inch in diameter and from 13–18 mils thick.

Samples compressed at pressures of 5000, 10 000, 20 000, 30 000, 40 000, 50 000, 100 000, 150 000, and 200 000 lb/in.^{2}, respectively, are shown in Fig. 1. The general appearance of the compressed samples ranged from a white disk to a completely translucent disk at around 125 000 lb/in.^{2} This phenomenon is not new and has been reported elsewhere.[1,2]

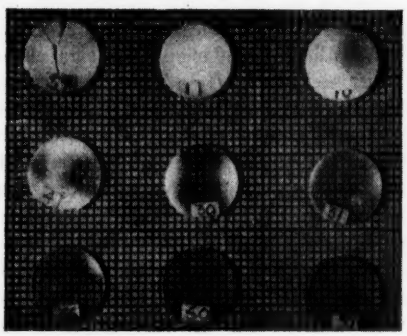


FIG. 1. KCl samples compressed at pressures of 5000, 10 000, 20 000, 30 000, 40 000, 50 000, 100 000, 150 000, and 200 000 lb/in.^{2}, respectively.

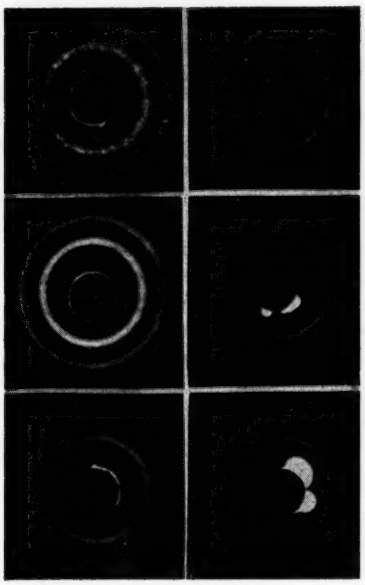


FIG. 2. X-ray diffraction patterns of compressed KCl samples.

Powder	5000 lb/in.^{2}
10 000 lb/in.^{2}	50 000 lb/in.^{2}
100 000 lb/in.^{2}	200 000 lb/in.^{2}

The compressed samples were then analyzed by the usual x-ray diffraction transmission methods using a North American Phillips Unit, Mo target tube operated at 50 kvp and 18 ma, zirconium filter, and a sample-to-film distance of 4.95 cm. The surface of the compressed samples was placed perpendicular to the undeviated x-ray beam.

The x-ray diffraction patterns of powdered KCl and of the compressed KCl samples are shown in Fig. 2. The powder sample gives a spotty pattern, indicating large grain size (60–80 microns). The patterns for the samples compressed at 5000 lb/in.^{2} and 10 000 lb/in.^{2} show fewer spots, indicating a decrease in grain size. The remaining patterns show that with additional pressure the rings become sharp and indicate a grain size of less than 5×10^{-3} mm.

Theoretically, as grain size continues to decrease to below 2×10^{-4} mm, the rings in the pattern begin to lose their sharpness and show signs of broadening, because such crystals lack the necessary resolving power to produce sharp rings.

Since this broadening of the rings was not evident to the eye, the patterns were further investigated with a microphotometer. The intensity films for the 50 000-lb/in.^{2} and 200 000-lb/in.^{2} diffraction patterns superimposed on each other, thus denoting no appreciable line broadening. Sproull[3] states that between 5×10^{-3} mm and 2×10^{-4} mm the x-ray pattern is insensitive to grain size. Therefore when a sample yields a sharp-ring pattern, one can only say that its grain size lies

within this range. The grain size of samples compressed from 10 000 to 200 000 lb/in.[2] appears to lie between 5×10^{-3} and 2×10^{-4} mm.

The finding that the respective Debye rings in the diffraction patterns remained the same distance from the undeviated x-ray beam for all of the compressed samples, indicated no change in crystal lattice structure as a result of the applied pressure. It appears that although these translucent compressed crystal samples resembled in appearance single crystals, they retained x-ray diffraction patterns characteristic of powder crystals. It may be that crystals prepared in this way can be used as phosphors in scintillation counters.

* Now at the Army Medical Research Laboratory, Fort Knox, Kentucky.

[1] J. G. Kereiakes, dissertation, University of Cincinnati, 1950 (unpublished).

[2] M. A. Ford and G. R. Wilkinson, J. Sci. Instr. **31**, 338 (1954).

[3] W. T. Sproull, *X-Rays in Practice* (McGraw-Hill Book Company, Inc., New York, 1946), p. 440.

Infrared Absorption of Liquid and Solid Hydrogen*

ELIZABETH J. ALLIN, W. F. J. HARE† AND R. E. MACDONALD‡

McLennan Laboratory, University of Toronto, Toronto, Canada

(Received February 28, 1955)

THE fundamental vibrational absorption of hydrogen, induced by intermolecular forces, has been studied over a wide range of gas pressures and at temperatures down to 80°K.[1–4] The resolution of the components of the band improves greatly as the temperature is lowered; for this reason, as well as to ascertain the effect of crystal structure on induced absorption, the investigation has been extended to liquid and solid hydrogen.

A fused quartz absorption cell, giving a path length of 8.4 mm, was suspended in a cryostat of special design. Temperatures in the range 10–22°K were maintained constant to ±0.1°K by controlling the boiling of liquid helium below the cell with a heating coil, and were measured by the vapor pressure of the hydrogen in the cell. A Perkin-Elmer 12B spectrometer with a LiF prism and a PbS detector was used. Figure 1 shows the fundamental absorption of liquid normal hydrogen at 18°K and of the solid at 11°K, along with the absorption of the compressed gas at 298°K with approximately the same density.

Several features of the absorption of the compressed gas are now well established. The Q_Q component [not well resolved under the experimental conditions of Fig. 1(a)] and various rotational lines with $\Delta J=\pm2$, of which $S(0)$ and $S(1)$ are the most prominent, have frequencies near those calculated for the free molecule at all gas densities; the intensity of these components is believed to arise chiefly from the quadrupole interactions between the absorbing molecule and the surrounding molecules.[5] The maxima Q_P and Q_R, on the other hand, change markedly with the density and the temperature; their separations from the band origin increase with increasing density and temperature, and the intensity of Q_P diminishes rapidly as the temperature is lowered. These components of the Q branch probably originate in close collisions in the region of overlap forces, and the relative kinetic energy of the surrounding molecules can apparently participate in the absorption process.[3]

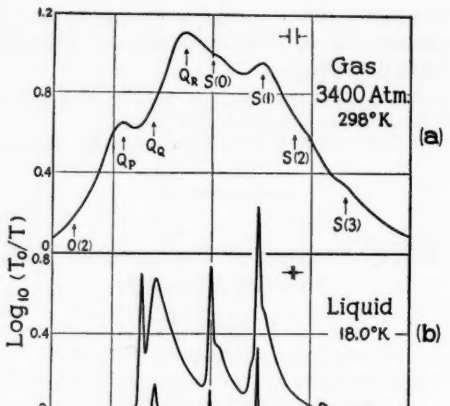


FIG. 1. The fundamental infrared absorption of hydrogen induced by intermolecular forces for (a) the compressed gas at a density of 1014 Amagat units, (b) the liquid at 817 Amagat, and (c) the solid at 972 Amagat.

In the spectrum of the solid, the sharp lines Q_Q, $S(0)$, and $S(1)$ in Fig. 1(c), can be ascribed to the quadrupole field effect. Confirmation of this view has been furnished by preliminary experiments on pure parahydrogen in which the disappearance of Q_Q, in accordance with the vanishing of the matrix element for the transition $J=0$ to $J'=0$,[5] has been established. Certain minor maxima can be identified from their frequencies as the double rotational transitions $2S(0)$, $S(0)+S(1)$, and $2S(1)$; in parahydrogen $2S(0)$ appears alone, as a sharp line with medium intensity. The weak maximum, $S''(1)$, belongs to the $S(1)$ group, but its origin is not clear.

The higher-frequency components of the S lines, labelled $S'(0)$ and $S'(1)$ in Fig. 1(c), have contours very similar to Q_R and are accordingly ascribed to overlap forces. The interaction with the kinetic energy of the surrounding molecules in the gas must here become an interaction with the vibrational energy of the

crystal lattice. The splitting off of the components, $S'(0)$ and $S'(1)$, is first noticeable at about 20°K and is more or less complete at 15°K; there is no change observed on passing the solidification point at 14°K, and very little on lowering the temperature to 4.2°K, the boiling point of liquid helium. If the characteristic appearance of the spectrum of the solid is due to the participation of the lattice modes in the absorption, it is clear that the crystalline structure persists locally into the liquid state several degrees above the fusion point.

Induced infrared absorption is caused by the *asymmetric* distortion of the absorbing molecule by intermolecular forces. In crystalline hydrogen asymmetric distortion can be produced both by the random orientations of the freely rotating molecules and by the vibrational motion of the crystal lattice. The experimental results indicate that the first type of distortion occurs for quadrupole interactions, and that both types are operative for overlap interactions.

The authors are indebted to Professor H. L. Welsh, Professor M. F. Crawford, Professor K. R. Atkins, and Professor A. C. Hollis Hallett for valuable discussions.

* This research was supported by a grant from the National Research Council of Canada.
† Holder of a research fellowship of the National Research Council of Canada.
‡ Holder of a scholarship of the Research Council of Ontario.
[1] Welsh, Crawford, and Locke, Phys. Rev. **76**, 580 (1949).
[2] Chisholm, McDonald, Crawford, and Welsh, Phys. Rev. **88**, 957 (1952).
[3] D. A. Chisholm and H. L. Welsh, Can. J. Phys. **32**, 391 (1954).
[4] Chisholm, Hare, and Welsh, J. phys. radium **15**, 511 (1954).
[5] J. van Kranendonk, dissertation, Amsterdam, 1952 (unpublished); J. van Kranendonk and R. Bird, Physica **17**, 153 (1952).

Canonical Invariance in Lagrangian Quantum Mechanics

W. K. Burton

Department of Natural Philosophy, University of Glasgow, Glasgow, Scotland

(Received January 17, 1955; revised manuscript received February 9, 1955)

Hugenholtz,[1] in a recent Letter to the Editor, has criticized Schwinger's[2] variational formalism in quantum mechanics by attempting to show that it is possible, by means of a unitary transformation on the dynamical variables, to add terms to the Lagrange function which are not in the form of time derivatives. He gives as an example $L''=-\frac{1}{2}(q^2\dot{q}+\dot{q}q^2)$ which arises from $U=\exp(\frac{1}{3}iq^3)$ and concludes that the new Lagrange function $L'=L+L''$ is 'essentially different' from the old one L.

This is not the case; the condition for canonical invariance is that the *variation* of the added term[2] rather than the added term itself should be a time derivative. If we evaluate $\delta L''$ we find

$$\begin{aligned}\delta L''&=-\tfrac{1}{2}(\delta q q\dot{q}+q\delta q\dot{q}+q^2\delta\dot{q}+\delta\dot{q}q^2+\dot{q}\delta qq+\dot{q}q\delta q)\\&=-(q\dot{q}+\dot{q}q)\delta q+\frac{d}{dt}(q^2)\cdot\delta q-\frac{d}{dt}(q^2\delta q)\qquad(1)\\&=-\frac{d}{dt}(q^2\delta q)=\frac{d}{dt}\delta(-\tfrac{1}{3}q^3),\end{aligned}$$

where we have used the commutativity of q with all its coefficients. Thus the new Lagrange function is equivalent to the old one.

The effect of making a general unitary transformation, not explicitly dependent on the time, can be seen as follows. First, every unitary transformation U can be written

$$U=\exp(iG),\qquad(2)$$

where G is Hermitian. Secondly, a unitary transformation of the dynamical variables at a time intermediate between the initial and final times has no effect on the transformation function:

$$\begin{aligned}&(\alpha_1't_1|\alpha_2''t_2)\\&=\int(\alpha_1't_1|\alpha't)d\alpha'(\alpha't|\alpha_2''t_2),\quad t_1>t>t_2\\&\rightarrow\int(\alpha_1't_1|U(t)|\alpha't)d\alpha'(\alpha't|U^{-1}(t)|\alpha_2''t_2)\qquad(3)\\&=(\alpha_1't_1|\alpha_2''t_2).\end{aligned}$$

This shows that (2) induces

$$(\alpha_1't_1|\alpha_2''t_2)\rightarrow(\alpha_1't_1|U^{-1}(t_1)U(t_2)|\alpha_2''t_2).\qquad(4)$$

Since the only change is at the times t_1 and t_2, Schwinger's operator form of Hamilton's principle is left unaffected. Thus if an admissible variation be made, the contribution to the variation of the transformation function arising from U is

$$-i(\alpha_1't_1|U^{-1}(t_1)[\delta G(t_1)-\delta G(t_2)]U(t_2)|\alpha_2''t_2).$$

But the same effect would be produced if we added $-(d/dt)\delta G$ to δL, since $\int dt(d/dt)\delta G=\delta G(t_1)-\delta G(t_2)$. Hence U induces

$$\delta L\rightarrow\delta L-(d/dt)\delta G.\qquad(5)$$

Hugenholtz's final point that c-number variations do not transform into c-number variations under arbitrary unitary transformations does not seem relevant. The point is more whether *admissible* variations transform into *admissible* variations. In all the cases which have been studied so far,[2] it can be seen that a variation consisting of the sum of two nonvanishing parts, one of which commutes and the other of which anticom-

mutes with the unvaried quantities, is an admissible variation. The commuting part produces no effect in Fermi systems, and the anticommuting part produces no effect in Bose systems. The full extent of the class of admissible variations is not yet known.

[1] N. M. Hugenholtz, Phys. Rev. **96**, 1158 (1954).
[2] J. Schwinger, Phys. Rev. **91**, 713 (1953).

Possibility of a Nuclear Radiation Detector by Means of Electroluminescent Phosphors

W. Low

Department of Physics, The Hebrew University, Jerusalem, Israel

(Received February 24, 1955)

EVER since Rutherford used scintillations of phosphorescent screens to count alpha particles, the scintillation counter has been used with great advantage as a nuclear radiation detector. Recently, in particular,[1,2] the combination of a photomultiplier with a phosphor like activated sodium iodide has served as a very efficient detector of α, β, and γ rays.

This note is to point out that the light output of scintillation counters using electroluminescent phosphors could possibly be increased by subjecting the phosphor to strong electric fields. (Strong electric fields have been used successfully with crystal counters like CdS and diamond.[3]) The application of dc or ac voltages of the order of 10^5 v/cm on certain powdered phosphors increases the luminescent brightness of ultraviolet-excited phosphors.[4–6] The increase of the brightness is a function of the applied voltage. Furthermore, the sudden application of an electric field or of light excitation (when the phosphor is subjected to the field) results in strong light flashes; this is called the Gudden-Pohl effect.[7] The electrons in the conduction band of a phosphor subject to high local fields are probably accelerated to energies which may cause further ionization, thus resulting in electron multiplication with perhaps local avalanche formation.[8,9] This process probably occurs only near the cathode[9] and is analogous to the processes found in glow gas discharges.

If an ionizing particle enters a phosphor which is subjected to very strong electric fields close to electrical breakdown, it will create local ionization and perhaps local avalanches along its path. These electrons on recombination will emit light flashes which can be amplified and recorded with a photomultiplier arrangement. Such a counter could possibly be made proportional by biasing the phosphor with different voltages.

We would like to suggest another possibility for the use of proper electroluminescent phosphors. Such phosphors show a threshold voltage above which light is emitted even without any external excitation. When the applied voltage increases, the light output also increases. The threshold voltage is considerably changed by irradiation with ultraviolet light, i.e., by the promotion of electrons into the conduction band. In such a case a momentary flash is observed. An ionizing particle may also give rise to a pulse of light from an electroluminescent phosphor.

[1] H. Kallmann, Natur. u. Tech. (July, 1947).
[2] R. Hofstadter, Phys. Rev. **75**, 796 (1949).
[3] R. Hofstadter, Proc. Inst. Radio Engrs. **38**, 726 (1950).
[4] Low, Steinberger, and Braun, J. Opt. Soc. Am. **44**, 504 (1954).
[5] G. Destriau, Electrochemical Society Meeting, May, 1954 (unpublished).
[6] Alexander, Low, and Steinberger (to be published). Observations have been made on ultraviolet-excited phosphors subjected to strong electric fields.
[7] B. Gudden and R. Pohl, Z. Physik **17**, 334 (1923).
[8] D. Curie, J. phys. radium **13**, 317 (1953).
[9] W. W. Piper and F. E. Williams, Phys. Rev. **81**, 151 (1952).

Cyclotron and Spin Resonance in Indium Antimonide

G. Dresselhaus, A. F. Kip, C. Kittel, and G. Wagoner

Department of Physics, University of California, Berkeley, California

(Received February 7, 1955)

CYCLOTRON resonance has been observed in a single crystal of indium antimonide at a frequency near 24 000 Mc/sec, using a light modulation technique with injected carriers, as described in our earlier publications.[1] One resonance line observed at low field strengths corresponds to an effective mass $m^*=(0.013 \pm 0.001)m$, and the resonance appears to be isotropic under rotation in a (100) plane, to within the experimental error. A recorder tracing of a typical run is shown in Fig. 1. We infer from the isotropic mass that the resonance is associated with conduction electrons. Pearson and Tanenbaum[2] have made magnetoresistance

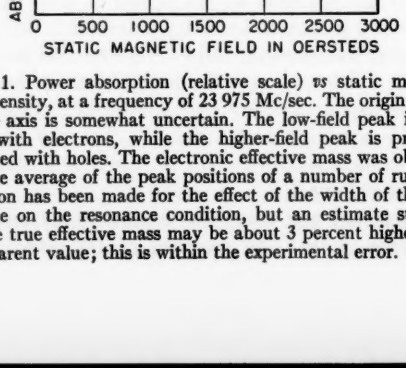


FIG. 1. Power absorption (relative scale) *vs* static magnetic field intensity, at a frequency of 23 975 Mc/sec. The origin for the vertical axis is somewhat uncertain. The low-field peak is associated with electrons, while the higher-field peak is probably associated with holes. The electronic effective mass was obtained from the average of the peak positions of a number of runs. No correction has been made for the effect of the width of the low-field line on the resonance condition, but an estimate suggests that the true effective mass may be about 3 percent higher than the apparent value; this is within the experimental error.

measurements on n-InSb which suggest strongly that the conduction band edge is spherical, while the work of Hatton and Rollin[3] suggests that the valence band is more complicated. Mobility and degeneracy studies by various workers had suggested a light effective mass for electrons in InSb, but not quite as light as we find at the band edge. The specimen giving the line shown in the figure is p-type with acceptor concentration estimated to be less than 5×10^{14} cm^{-3}. At 4°K the resonance can barely be resolved, but at 2.2°K the line is clearly resolved, with a relaxation time of approximately 1×10^{-11} sec. An n-type specimen with $N_d\sim10^{15}$ cm^{-3}, $N_a\sim10^{13}$–10^{14} cm^{-3}, gave at 2.2°K a relaxation time of about 0.7×10^{-11} sec.

We have made preliminary observations on two resonance lines with heavier masses, $m^*\sim0.18m$ and $m^*>1.2m$; the lines are somewhat anisotropic, and we suggest tentatively that they may be associated with holes. The study of these lines is continuing.

An electron spin resonance line has also been observed in both crystals at a g value of 2.11 ± 0.03.

We are particularly indebted to H. J. Hrostowski of the Bell Telephone Laboratories for the provision of the InSb crystals. The high purity of these crystals made possible the work reported here. We are indebted also to R. C. Fletcher and C. Herring for assistance and information. This work has been supported in part by the Office of Naval Research and the U. S. Signal Corps.

[1] A. F. Kip, Physica (to be published); Dresselhaus, Kip, and Kittel, Phys. Rev. **98**, 368 (1955).
[2] G. L. Pearson and M. Tanenbaum, Phys. Rev. **90**, 153 (1953).
[3] J. Hatton and B. V. Rollin, Proc. Phys. Soc. (London) **A67**, 385 (1954).

PHYSICAL REVIEW VOLUME 98, NUMBER 2 APRIL 15, 1955

Proceedings of the American Physical Society

MINUTES OF THE SEVENTH ANNUAL GASEOUS ELECTRONICS CONFERENCE (HELD UNDER THE JOINT SPONSORSHIP OF THE DIVISION OF ELECTRON PHYSICS, AMERICAN PHYSICAL SOCIETY, AND THE COLLEGE OF ENGINEERING, NEW YORK UNIVERSITY) AT NEW YORK, NEW YORK, ON OCTOBER 14–16, 1954

THE Seventh Annual Conference on Gaseous Electronics was held in Vanderbilt Hall on the Washington Square Campus of New York University on October 14, 15, and 16, 1954. There were 244 registrants. A symposium of three invited papers consisted of "Mechanism of Uniform Field Breakdown," by F. Llewellyn Jones, University Collete of Swansea; "Formative Time Lags of Uniform Field Breakdown," by L. H. Fisher, New York University; and "Secondary Processes and the Mechanisms of Spark Breakdown," by L. B. Loeb, University of California, Berkeley. A fourth invited paper, "Effect of Electrons on Spectral Lines Emitted in a Discharge," was by H. Margenau of Yale University. Abstracts of some of the contributed papers are printed hereunder.

The banquet of the Conference was held on Friday evening, October 15. T. E. Allibone (Associated Electrical Industries of Great Britain) and W. Finkelnburg (Siemens-Schuckert Werke, Erlangen, Germany) were the after-dinner speakers, and the theme of each was the state of Gaseous Electronics research in his country.

The Conference Committee for the coming year consists of W. P. Allis (*Chairman*), J. D. Cobine, L. H. Fisher, H. Margenau, A. V. Phelps, and D. J. Rose.

LEON H. FISHER For the Conference Committee

A1. Negative Ion Formation Using Monoenergetic Electrons. W. M. HICKAM AND R. E. FOX. *Westinghouse Research Laboratories.*—Ion formation involving electron attachment at low energies was investigated with monoenergetic electrons as obtained by the retarding potential difference (RPD) method.[1] Initial studies with this method reveal that electron attachment in some cases may occur only over an extremely narrow energy range. For such cases it becomes increasingly important that monoenergetic electrons be used in order to obtain the true excitation curve. For example, in the case of the formation of SF_6^-, the large cross section is found to occur only for electrons whose energy is less than 0.1 ev. The dissociative attachment process which yields SF_5^- also has a maximum cross section for electron energies of less than 0.1 ev. In this case, however, the negative ion formation as a function of electron energy decreases more slowly, going to zero at about 2 ev. The use of monoenergetic electrons can account for the wide differences between these results and those previously obtained[2] for SF_6. It will be demonstrated how the extremely narrow capture process can yield information concerning the behavior of the electron beam at these low energies.

[1] Fox, Hickam, Kjeldaas, and Grove, Phys. Rev. **84**, 859 (1951).
[2] A. G. Ahearn and N. B. Hannay, J. Chem. Phys. **21**, 119 (1953).

A2. Rotational Excitation by Slow Electrons.* E. GERJUOY AND S. STEIN, *University of Pittsburgh and Westinghouse Research Laboratories.*—Theoretical cross sections have been developed for the rotational excitation of homonuclear molecules by slow electrons,[1] and the results applied to calculation of the fractional energy loss per collision λ. In N_2, at energies below the vibrational threshold, theoretical losses are approximately twice experimental[2] but many times larger than $2m/M$. In H_2 the theoretical losses are not more than 2.5 ($2m/M$), and, except at the lowest energies studied (ν 0.1 ev), are significantly smaller than observed.[2] It would be desirable to have more direct experimental evidence of rotational excitation. For this reason we have calculated λ at 77°K in pure para-hydrogen and in the equilibrium mixture of ortho- and para-hydrogen. At electron energies $\sim$0.075 ev,

the two λ's should differ by about 50 percent. Similarly, because of the altered rotational distribution, λ for deuterium differs from λ for H_2. Such differences, if observed, could hardly be accounted for on any other basis than rotational excitation.

* This research was supported in part by the U. S. Air Force.
[1] Stein, Gerjuoy, and Holstein, Phys. Rev. **93**, 934 (1954).
[2] R. W. Crompton and D. J. Sutton, Proc. Roy. Soc. (London) **A215**, 467 (1952).

A3. Redetermination of the Ionization Coefficient η for Hydrogen. D. J. ROSE, *Bell Telephone Laboratories.*—Measurements of the ionization coefficient for hydrogen have been reported by Ayers (1923) and Hale (1938). These data are inconsistent. The experiment has therefore been repeated, using the best available vacuum technique. The hydrogen was prepared either from spectroscopically pure one-liter flasks, or from the decomposition of uranium hydride which was prepared in vacuum from the pure gas and reactor grade uranium. The plane parallel electrodes were 5-cm disks of molybdenum, and could be adjusted to pre-set separations within 10 microns error. Values of the ionization coefficient η were obtained in the range $20<E/p<1000$ volts/cm×mm Hg.

A4. Studies of Anode Pre-Breakdown Phenomena in Thyratrons. SEYMOUR GOLDBERG, *Edgerton, Germeshausen & Grier, Inc.*—The transfer of gaseous conduction from a region containing a plasma to a space which is separated from this by a system of apertures and baffles and which contains an anode has been investigated. The dc anode voltage-anode current characteristics of the system have been obtained and indicate that the anode behaves much the same as a Langmuir probe, permitting the evaluation of electron temperature and space potential. Reasonable values for these parameters are obtained. It is found that when the potential is raised above that of the plasma, the current increases beyond the saturation value, probably due to an expanding electron sheath about the probe. As the potential is raised further, considerable ionization is apparent in the anode space and a plasma-like boundary is observed to protrude through the apertures partitioning the regions. The effect of ionization in the anode region is to increase potentials in the connecting space, which further expands the electron sheath. At some critical current to the anode this boundary abruptly increases and breakdown occurs.

A5. Diffusion of Ions and Electrons in an Electric Field. ROBERT N. VARNEY, *Washington University.*—The partial differential equation for diffusion has been set up to include a drift velocity term and has then been solved explicitly by standard Fourier integral procedures. The results are then applied to three ionic problems. The first is the diffusion of an ion avalanche cloud opposite to the field against an anode surface. The second is the analysis of the approach of a positive ion avalanche to a collecting cathode under the influence of both drift and diffusion. The third is the analysis of back-diffusion of electrons after escape from a cathode to the cathode surface. In all cases, a variable of controlling significance is the function vx/D, where v is the drift velocity, D the diffusion coefficient, and x a distance of travel. The function vx/D shows diffusion effects to be negligible if it is appreciably less than unity.

A6. Ionization Accompanying the Long-Lived Nitrogen Afterglow.* W. B. KUNKEL AND A. L. GARDNER, *University of California, Berkeley.*—Extensive electron density measurements have been carried out in active nitrogen by means of the microwave method. It was found that the electron density depends strongly on the amount of contamination in the nitrogen, being highest for an admixture of 0.1 percent to 0.2 percent of oxygen and reaching a maximum at a time of the order of 100 milliseconds after the discharge excitation, in striking agreement with the emission of the visible spectrum. It is concluded that the oxygen, being totally dissociated in the discharge, does not attach electrons but serves to form nitric oxide which has a sufficiently low ionization potential to be ionized by the active particles. Since electron removal by ambipolar diffusion and recombination is rapid, two-body ionizing collisions are required to maintain the high free electron density level, ($\sim 10^9$ cm^{-3}). Metastable nitrogen molecules of 9.6 ev, presumably in a singlet state, must therefore be continually formed out of the parent ground state atoms. It seems possible that the metastables are also needed as intermediate states in the afterglow reaction and that the presence of the electrons is essential for the multiplicity change required for the light emission as suggested by Nicholls.[1]

* Work supported by Wright Air Development Center.
[1] R. W. Nicholls, J. Chem. Phys. **20**, 1040 (1952).

A7. Diffusion and Volume Loss Coefficients of Helium Metastable Atoms and Molecules. A. V. PHELPS, *Westinghouse Research Laboratories.*—Diffusion and volume destruction coefficients have been determined for the metastable states of the helium atom,[1] 2^1S and 2^3S, and molecule, $2^3\Sigma$, from measurements of the lifetimes of these metastables in very pure helium. The product of the diffusion coefficient and the neutral atom density for each of the metastable states is 1.53×10^{19} (cm²/sec) (atom/cc) at 300°K to within the respective experimental errors. The volume loss for the 2^1S metastables is linear with the helium pressure and suggests destruction by collision induced radiation with an average cross section of 4×10^{-20} cm² at 300°K. The triplet metastable atoms are destroyed in three-body collisions[1] with a combination coefficient of 2.5×10^{-34} cc²/atom²-sec at 300°K. It has not been possible to observe a loss of molecular metastables which increases with helium pressure because of destruction in collisions with slow electrons or other metastables. These high-pressure experiments show that the natural lifetime of the molecular metastables is at least 0.05 second and that the volume destruction is at least one hundred times smaller than that for the atomic metastables.

[1] A. V. Phelps and J. P. Molnar, Phys. Rev. **82**, 1202 (1953).

A8. The Effect of Temperature on the Lifetime of Metastable Mercury Molecules. C. G. MAITLAND AND A. O. MCCOUBREY, *Westinghouse Research Laboratories.*—The band radiation from mercury vapor in an afterglow excited by the 2537 A resonance line has been accounted for in terms of metastable mercury molecules $Hg_2(^3O_u^-)$.[1] These molecules are destroyed by spontaneous radiation, three-body collisions, and upon diffusing to the walls. In the present work, the effect of temperature on these processes has been measured. The experimental method consists of measuring the decay time of the afterglow over a range of mercury vapor densities at several constant vapor temperatures. As the vapor temperature is varied from 380°K to 580°K, the results obtained are as follows: (1) the spontaneous radiation lifetime is found to decrease from 63 milliseconds to 19 milliseconds; (2) the diffusion coefficient[2] follows a $T^{\frac{1}{2}}$ variation within the experimental error (±4 percent); (3) the three-body collision induced radiation rate, $C=29\times10^{-33}$ (atoms/cc)$^{-2}$ sec^{-1}, is unchanged.

[1] A. O. McCoubrey, Phys. Rev. **93**, 1249 (1954).
[2] A. O. McCoubrey and C. G. Matland, Phys. Rev. **96**, 832(A) (1954).

B1. Concerning Oscillations of a Plasma with a Skew Velocity Distribution Transverse to a Static Magnetic Field.* J. E. DRUMMOND AND L. WILCOX, *Sylvania Electric Products, Inc.*—The present paper extends the analysis of plasma oscillations in a magnetic field developed by E. P. Gross.[1]

Assuming an isotropic steady state electron velocity distribution, Gross found gaps in the spectrum of possible traveling waves at integral multiples of the cyclotron frequency for electrons. The effect of a skewed steady velocity distribution on the dispersion relation has been studied and it is found that traveling wave solutions may exist at these "gap frequencies." A condition has been derived for traveling wave solutions in terms of parameters of the distribution function. It is pointed out that a small constant transverse electric field may be included in Gross' analysis and is equivalent in first approximation to the case treated.

* Work carried out under U. S. Signal Corps Contract.
[1] E. P. Gross, Phys. Rev. **82**, 232 (1951).

B2. Plasma Fluctuations in Crossed Electric and Magnetic Fields. HUGH W. BATTEN, *University of Michigan.*—An experimental study of the spontaneous, high-intensity electric fluctuations associated with a gas discharge in a magnetic field was undertaken. Probes were introduced into the plasma region of an especially designed gas diode, and the voltages on the probes were observed. Data were taken to cross-correlate the electric fluctuations on two neighboring probes in the discharge. These fluctuations on the two probes were similar except that those on the "downstream" probe were delayed in time with respect to those on the "upstream" probe. Interpreting this delay as the time of propagation from one probe to the other, the velocity and direction of propagation could be determined. The results show that the fluctuations are hundreds of volts in amplitude, propagate approximately in the direction of electron drift, and have a speed somewhat less than the Lorentz drift velocity. There is experimental evidence to suggest that these fluctuations grow as they propagate.

B3. Instantaneous Probe Measurements of Traveling Striations in an Argon Glow Discharge.* N. L. OLESON, *U. S. Naval Postgraduate School.*—Further measurements have been made of the instantaneous probe currents obtained from the positive column of an Argon glow discharge in which moving striations are present. Simultaneously, the light from the neighborhood of the probe is scanned by a photomultiplier and probe current, and light intensity are presented on the screen of a dual beam oscilloscope. The discharge tube has a diameter of 2 centimeters and observations have been made with pressures from one-half millimeter to twelve millimeters and currents from 20 to 40 milliamperes. The probe currents are characterized by sharp spikes where the current becomes electronic for a very short interval. The location of these spikes with respect to maximum light intensity depends upon probe position, gas pressure, and current. As the probe is made more positive with respect to the anode, broader electron current peaks appear and the average current to the probe becomes electronic.

* Supported by the Office of Naval Research.

B5. Some Aspects of Plasma Resonance.* K. S. W. CHAMPION† AND S. C. BROWN, *M.I.T.*—The impedance (ρ) to high-frequency electromagnetic waves of any region of a plasma, divided by the wavelength, (ρ/λ) is plotted as a function of $n\lambda^2$ (where n is the electron density) and $p\lambda$ (where p is the pressure). The resulting surface is of particular interest in the vicinity of plasma resonance and is important for two reasons. One is that it gives a plot of the impedance any plasma presents to the transmission of microwaves. It not only confirms the measurements of Goldstein[1] and Makinson of the microwave impedance of dc plasma, but also indicates immediately the range over which plasma resonance properties can be used as a "microwave probe" to measure electron densities. Experimentally this technique of density measurement has been used with high-frequency discharges for the first time. The second reason for the importance of the impedance plot is that it enables the properties of microwave discharges to be predicted. In particular, more extensive calculations have been made which enable the appearance of any microwave discharge to be predicted quantitatively, provided its diameter exceeds four times its height.

* This work was supported in part by the Signal Corps, the Office of Scientific Research, Air Research and Development Command, and the Office of Naval Research.
† Now at Tufts College.
[1] L. Goldstein and N. L. Cohen, Elec. Commun. **28**, No. 4, 305 (1951).

B6. Microwave Study of Positive Ion Collection by Probes. GEORGE J. SCHULZ* AND SANBORN C. BROWN, *M.I.T.*—The positive ion saturation region of Langmuir probes may be a convenient tool for measuring positive ion densities in a plasma when the probe does not exhibit a sharp electron saturation (e.g., at higher pressure). The theory of the positive ion saturation region is modified to include the drift current of positive ions at the edge of the sheath, $J_d = ne \times (kT/M)^{\frac{1}{2}}$ where J_d is the directed current density at the sheath edge, k is Boltzmann's constant, T_- is the electron temperature, n is the electron or positive ion density, M is the mass of the ion, and e is the electronic charge. This modification explains the large positive ion saturation currents at low pressure reported by many authors.[1] The current collected by a cylindrical probe is determined not only by J_d, but also by the collection efficiency of the probe. When no collisions occur in the sheath, ions describe orbital motions around the probe. When approximately one collision occurs on the average, most ions will lose energy and become trapped in the sheath and eventually reach the probe. The theory is equally applicable to single probes and to double probes.[2] These considerations were confirmed by taking probe curves with single and double probes in a microwave gas discharge and using microwave methods for an independent measurement of electron density.

* Now at Westinghouse Research Laboratories, East Pittsburgh, Pennsylvania.
[1] R. M. Howe, J. Appl. Phys. **24**, 881 (1953).
[2] E. O. Johnson and L. Malter, Phys. Rev. **80**, 58 (1950).

B7. Cumulative Ionization Processes in the Helium Positive Column. L. S. FROST AND A. V. PHELPS, *Westinghouse Research Laboratories.*—Preliminary measurements of important parameters of a helium positive column have been made in order to investigate cumulative ionization. Electron density is measured by microwave techniques, the density of singlet and triplet He metastable atoms by optical absorption, and the axial electric field and the average electron energy by fine wire probes. The total ionization rate is calculated by the diffusion loss equation. The rate of direct ionization of normal He atoms by high-energy electrons is calculated from published microwave breakdown data for He,[1] assuming negligible electron-electron interaction. The difference between the rates of total and direct ionization is the rate of cumulative ionization. This is compared with the ionization rate due to ionizing collisions between pairs of triplet metastable atoms,[2] calculated using the cross section obtained from afterglow studies. Results of this comparison together with evidence on the relative importance of other possible cumulative ionization processes will be presented.

[1] F. H. Reder and S. C. Brown, Phys. Rev. **95**, 885 (1954).
[2] M. A. Biondi, Phys. Rev. **88**, 660 (1953).

C1. Effect of Mean Free Path Variation on High-Frequency Breakdown. I. A. MACLENNAN AND A. D. MACDONALD, *Dalhousie University.*—The effect of the electron mean free path on the high-frequency breakdown field of gases is investigated in a neon-argon mixture. The mixture is predominately neon, the argon being added so that excited neon atoms may collide with argon atoms to produce

ionization (Penning effect). Assuming that electrons are produced by field ionization, and lost by diffusion, the Boltzmann transport equation is solved approximately for various elastic collision cross sections, Q. The theory shows that the variation in Q for low-energy electrons considerably changes the breakdown condition, and suggests that information about the elastic collision cross section in some gases might be obtained from the breakdown measurements.

C2. Departure from Paschen's Law of Breakdown in Gases. W. S. BOYLE AND P. KISLIUK, *Bell Telephone Laboratories.*—The failure of Paschen's law of electrical breakdown in gases at both high pressures and extremely small electrode separations is explained by a single breakdown mechanism. The breakdown field in these two regions is sufficiently great to draw measurable field emission currents from the cathode, which produces a relatively small number of ions. The space-charge field of these ions is great enough to increase appreciably the field emission current even when the ratio of ion current to electron current is less than one percent. As more ionic space charge is produced each ion becomes more effective in enhancing the electron current until the breakdown condition is attained. An expression is derived for the yield of electrons per positive ion as a function of the applied field. This dependence is shown to be in agreement with experimental breakdown voltages at high pressure. It is also shown that the conditions leading to breakdowns at extremely short distances are appropriate for the same processes to take place, thus explaining breakdowns below the "minimum sparking potential."

C3. Electrical Breakdown in High Vacuum. P. KISLIUK, W. S. BOYLE, AND L. H. GERMER, *Bell Telephone Laboratories.*—Currents preceding breakdown have been measured between closely spaced tungsten electrodes in high vacuum. It is found that field emission currents sufficient to evaporate anode metal flow before breakdown. These currents follow the Fowler-Nordheim equation when field magnification due to surface irregularities on the cathode is taken into account. The field magnification is a function of distance at electrode separations less than 4×10^{-4} cm. Explanation of the observed breakdown at low voltage and small spacing requires an unusually high yield of electrons at the cathode per ion formed in the gap. Furthermore there is no measurable direct enhancement of the current by ionization even at higher voltages. The high electron yield must therefore exist over the entire observed range of breakdown voltages. This high yield is satisfactorily accounted for by the increase in field emission due to the positive ion space charge, which in turn increases the positive ion current density until there is breakdown. Breakdown occurs when the field-emission current is increased by only 65 percent. This condition is reached with the ion current density much smaller than the electron current density.

C4. Sparking Potential and Molecular Structure of Hydrocarbon Gases. R. W. CROWE AND J. C. DEVINS, *General Electric Research Laboratory.*—In order to calculate the Townsend α for a gas from the energy distribution of conducting electrons, one must have information about certain physical properties of the gas molecules, such as their cross sections for elastic scattering and energy dependence of their cross sections for excitation and ionization. Although such information is rarely available for molecular gases, a semi-empirical expression for α may be developed by the introduction of appropriate approximations in the theoretical treatment. For example, if the total inelastic (excitation+ionization) cross section can be approximated by $\sigma=\sigma_0(\epsilon-\epsilon_i)$, where ϵ_i is the excitation threshold, then it is possible to show from the derivations of Smit[1] that α should be of the form,

$$\alpha/P\cong A\,(E/P)\,\exp[-K(\sigma_0\theta)^{\frac{1}{2}}(\epsilon_i-\epsilon_i)^{\frac{1}{2}}(P/E)],$$

where θ is the elastic cross section of the molecule, and ϵ_i is its ionization energy. The purpose of this investigation was to determine the influence of small variations in molecular structure upon the composite cross section $(\sigma_0\theta)^{\frac{1}{2}}$. This was accomplished by measuring the sparking potentials of a large number of hydrocarbon gases as functions of $P\delta$ and comparing the results with the expression derived from a combination of the above equation with the Townsend criterion for breakdown. A simple correlation between $(\sigma_0\theta)^{\frac{1}{2}}$ and the number of the various chemical bonds in the hydrocarbon molecule was found. An attempt is made to interpret such a correlation in terms of electronic structure.

[1] J. A. Smit, Physica 3, 543 (1936).

C5. Breakdown Processes in Nitrogen, Oxygen, and Mixtures.* ELSA L. HUBER, *University of California, Berkeley.*—The studies of corona mechanisms in coaxial cylindrical geometry with central anode using α particle triggering, initiated by Colli and Facchini and by Lauer in inert gases, have been extended to pure nitrogen, pure oxygen, and mixtures of these gases. In pure nitrogen, from 25 to 600 mm pressure, the predominant mechanism is secondary liberation of electrons from the cathode by positive ion impact, with γ_i varying between 1.4×10^{-3} and 0.6×10^{-3}. The ion is the N_4^+ observed by Varney. Addition of the order of 1 percent of oxygen reduces γ_i by a factor of $\frac{1}{2}$ and gives a weak photoelectric γ_p at the cathode, both effects resulting in part from action of oxygen on the cathode. Photoionization *in the gas* also appears and with 5 percent oxygen the self-sustaining corona consists to a large extent of Geiger counter like pulses propagating along the wire. With the use of transverse and longitudinal α particle trajectories relative to the axis, the velocity of propagation of the discharge down the wire has been measured for various pressures in air, and for constant pressure and potential as a function of oxygen concentration. Velocities ranged between 10^8 and 10^7 cm/sec. In pure oxygen above 200 mm pressure, the absorption of photoionizing radiation is so intense that instead of spreading along the wire, the discharge propagates outward in the form of radial streamers.

* This work was supported at various stages by both the Office of Naval Research and the National Science Foundation.

C6. Measurement of the Current During the Formative Time Lag of Sparks in Uniform Fields in Air.* H. W. BANDEL, *University of California, Berkeley.*—The current in a parallel plane gap during the long formative time lags of sparks in air has been measured from 10^{-6} amp (a few μsec after application of the voltage) up to 10^{-2} amp (just before breakdown), for time lags between 10 and 100 μsec. Calculations for the increase of $e^{\int\alpha dx}$ due to field distortion by positive ions are in good agreement with an observed rapid increase of current shortly before breakdown. The Townsend discharge is shown to be spread over the whole electrode surface in contrast to the filamentary nature of the final spark. The photosensitivity of the cathode surface has been found to decrease following collection of positive ions. The conditioning of the electrodes observed by Fisher and Bederson is believed to be due to such a decrease in γp. Some gain has been made toward a theoretical solution for the buildup of the Townsend discharge. Working with an integral equation and taking due care with integration limits, has established the range of validity of the existing solution and yielded another approximate solution; from these, exact solutions for the limiting cases of only one γ are obtained. There is over-all agreement between theory and experiment within limits of error except for an observed delay in the initial current rise. To explain this requires the consideration of mechanisms involving

creation and transmission of active photons, which could cause a delay of the order of an electron crossing time.

* Work supported by Office of Naval Research, National Science Foundation, and Research Corporation.

C7. Investigation of Buildup Processes in an Argon Discharge. M. MENES, *Westinghouse Research Laboratories.*—Formative time lag studies indicate that the spark breakdown of atmospheric pressure argon is preceded by a Townsend buildup involving a delayed photon gamma mechanism.[1] Such a delayed photon emission has been found in argon counter discharges, and has been attributed to a long-lived molecular state.[2] The present experiment was intended to investigate further the processes active in the uniform field breakdown of argon. The experiment consists in recording oscillographically the buildup of current and of the light emission in a gap after application of voltage to the gap. The observed current buildup is exponential except for the final stage. The results have been analyzed on the basis of a delayed photon emission as proposed by Colli; they are in fair agreement with Colli's data at atmospheric pressure, but deviate appreciably at lower pressures. The cause of this discrepancy is not known and further experimental work is in progress. The final stages of the buildup appear to be controlled by space-charge field distortion, but no streamer mechanism, such as described for air, has been found in argon.

[1] G. A. Kachikas and L. H. Fisher, Phys. Rev. 91, 775 (1953).
[2] L. Colli, Phys. Rev. 95, 892 (1954).

E1. Ejection of Electrons from Contaminated Metals by Positive Ions.—HOMER D. HAGSTRUM, *Bell Telephone Laboratories.*—The studies of electron ejection by noble gas ions from metals which are either atomically clean or have a monolayer of gas upon them have been extended to contaminated metal surfaces. The ions used were He^+ and Ne^+. The metal targets, tungsten and molybdenum, were studied in four conditions: (1) as put into the experimental tube; (2) after heating to "baking temperatures"; (3) after heating to 1330°K; and (4) atomically clean. Measurements were made of total electron yield γ_i for ions of kinetic energy 10 to 1000 ev and of ejected electron energy distributions at 40 and 1000 ev. The results for the target in conditions (1) and (2) (which we call contaminated) are much the same. As judged from the electron ejection measurements, the target in condition (3) looks much like one covered with a monolayer of nitrogen. Whereas γ_i is always less than 0.3 and roughly independent of ion energy for a clean surface, for a contaminated surface it depends strongly on ion energy and may reach values approaching unity for 1000 ev ions. Electrons ejected from the contaminated surface are slower than those ejected from a clean surface. It is evident that ordinary baking cleans the surface of a metal very little and that much of the published work on electron ejection by ions was carried out with contaminated surfaces.

E2. Measurements and Collision Theory of High Vacuum Sputtering. FRANK KEYWELL, *Bell Telephone Laboratories.*—A Philips Ionization Gauge discharge has served as an ion source to measure absolute sputtering ratios due to ions in the energy range 400–6100 volts for the gas-metal combinations: Ag−Kr, Ag−A, Ag−Ne, Ag−He, Cu−Kr, Cu−A, Pb−A, Pb−He. The data has been interpreted by use of classical hard collision theory as described by Seitz.[1] This leads to a sputtering ratio formula which is fitted to the data by means of an absorption parameter, α, introduced to take into account the effect of recoil atoms being displaced away from the surface or being stopped before reaching the surface. Since the phenomenon of sputtering is a process of momentum transfer, it should be possible to recognize an effect on the sputtering ratio due to ions which rebound from the surface on the first collision. The data from this experiment has been treated by an "equivalent ion shift," which displaces the various curves from their relative positions to positions which correspond to equal average momentum transfer at the first collision. This tends to merge the data but there is some scant evidence that the ion rebound effect causes the shifted curves for the light ions to be low. The effect could be given a better test by means of sputtering measurements at threshold energy or energy ~15 kv on a gas-metal combination such as Cu−Kr, Cu−A, having about the same average energy transfer per collision but a metal atom of mass intermediate to the gaseous ions.

[1] F. Seitz, Disc. Faraday Soc. 271 (1949).

F1. Absorption of Radiation in the Vacuum Ultraviolet by *N* Atoms Produced in a Discharge Plasma in Nitrogen.* A. W. EHLER AND G. L. WEISSLER, *University of Southern California.*—Absorption of electromagnetic radiation [in the vacuum ultraviolet region of the spectrum by a Phillips Ionization Gauge (PIG) type discharge plasma in N_2] was investigated. Light from a spark discharge source followed a path through the PIG plasma and was analyzed by means of a vacuum spectrograph. Two photographic plate exposures were made of the source emission spectrum: one with the PIG discharge "on" and the other with the discharge "off." With the densitometer traces of these two exposures the composite absorption cross section of the PIG plasma was determined between 800 A and 400 A, and found to be about 17×10^{-18} cm². Then the absorption cross sections and relative concentrations of the various plasma constituents were considered in the light of available information in the literature. Evidence was presented which indicates that the significant absorbing component in these measurements was atomic nitrogen. The absorption cross section of atomic nitrogen as found by this analysis agrees well within the error limit of this experiment with the theoretical value calculated by Bates.[1]

* Sponsored by the Geophysics Research Directorate of the Air Force Cambridge Research Center.
[1] D. R. Bates and M. J. Seaton, Monthly Notices Roy Astron. Soc. 109, 698 (1949).

F2. Photoionization Cross Sections of CO_2, A, H_2, and H_2O.* N. WAINFAN, W. C. WALKER, AND G. L. WEISSLER, *University of Southern California.*—The photoionization cross sections of several gases were determined in the wavelength range from 473 A to 1100 A.[1] Simultaneous measurements of the total absorption cross sections and ionization efficiencies were made in order to compute these ionization cross sections. The total absorption cross sections obtained are in agreement with those reported previously.[2–4] In addition, the position of the ionization onset in each case agrees with the value of the first ionization limit obtained by other methods. For CO_2, the photoionization cross sections were found to fluctuate rapidly in the region from the ionization limit to 690 A. This fluctuation is presumably due to the presence of strong molecular absorption bands known to fall in this region. At wavelengths below 690 A, the cross sections of CO_2 are between 30×10^{-18} and 40×10^{-18} cm². The ionization cross sections of A in the region from the ionization onset to 473 A were found to be between 10×10^{-18} cm² and 25×10^{-18} cm². A value of 17×10^{-18} cm² was obtained at 770 A, that is, just beyond the onset. This value is somewhat lower than the theoretical estimate for argon made by Dalgarno[5] of 30×10^{-18} cm². A preliminary study of H_2 in the region near the ionization limit yielded ionization cross sections of the order of 8×10^{-18} cm², which verifies the estimate of the magnitude of the ionization continuum made by Lee and Weissler.[3] For H_2O vapor, preliminary results give ionization cross sections ranging from about 5×10^{-18} cm² near the ionization limit to 20×10^{-18} cm² at about 650 A. At present, work is in progress

on these and other gases as well as on various problems associated with the mechanism of photoionization.

* The aid of the Office of Naval Research is gratefully acknowledged.
[1] Wainfan, Walker, and Weissler (to be published).
[2] H. Sun (private communication, 1954).
[3] Po Lee and G. L. Weissler, Astrophys. J. 115 (May, 1952).
[4] Astoin, Johnn, and Vodar, Compt. rend. 237 (1953).
[5] A. Dalgarno, Proc. Phys. Soc. (London) 66, 392, 663–667 (1953).

F4. Further Results on the Photoelectric Yields of Metals in the Vacuum Ultraviolet.* W. C. WALKER, N. WAINFAN, AND G. L. WEISSLER, *University of Southern California.*—The absolute photoelectric yields of several metals subjected to various surface treatments were measured in the wavelength region from 1400 to 473 A using techniques previously described.[1] Photocathodes of Ni and Cu were studied: (1) in the untreated state, (2) after being heated in vacuum, (3) after being heated in an atmosphere of oxygen at reduced pressures, (4) after being heated in H_2, and (5) while heated red hot in vacuum. Hot samples of Pt and Au were also investigated. For each metal tested, the yields from the hot sample were found to be a factor of two lower than those obtained from the same sample at room temperature and a pressure of 5×10^{-5} mm of Hg. It was also found that at wavelengths longer than about 1000 A the photoyields of each of the metals decreases rapidly. In the case of Ni, oxygen treatment resulted in an increased yield due to the formation of oxides which could subsequently be removed by heating in hydrogen. The photoyields of copper were found to be insensitive to oxygen treatment and only slightly affected by moderate heating in hydrogen. Both these facts indicate that for copper, the results should be considered characteristic of a copper oxide rather than a metallic copper surface. Further work is now in progress on problems concerning the energy distribution of the photoelectrons and the effect of oxide formation on the yields of some other metals.

* The aid of the Office of Ordnance Research, Department of the Army, is gratefully acknowledged.
[1] Wainfan, Walker, and Weissler, J. Appl. Phys. 24, 1318–1321 (1953).

F5. Motion and Spectrum of Arc Cathode Spot in a Magnetic Field. R. M. ST. JOHN* AND J. G. WINANS, *University of Wisconsin.*—The velocity of the cathode spot of a mercury arc in a transverse magnetic field[1] has been measured for magnetic field strengths between 0 and 20 700 oersteds. The approximate doubling of retrograde velocity at about 11 000 to 15 000 oersteds was followed by an additional rapid rise of velocity at about 15 000 oersteds. Spectra of the arc showed Hg II and Hg III lines at the stronger magnetic fields. Radiation from the cathode spot showed mercury lines and a continuous spectrum which is especially intense at the lines. Some Hg lines are broadened asymmetrically and others symmetrically. The cathode spot radiation appears to come from excited gas outside the cathode surface. An arc mechanism previously proposed[1] is extended to explain the rapid velocity rises with increasing magnetic field strength by associating them with Hg^{++} and Hg^{+++} ions.

* Now at University of Oklahoma.
[1] R. M. St. John and J. G. Winans, Phys. Rev. 94, 1097 (1954).

F6. Anchored Mercury Arc Cathode Spot. CHARLES G. SMITH, *Raytheon Manufacturing Company.*—A mercury arc was anchored to a polished molybdenum cylinder (axis vertical, diameter 2 cm) projecting about 3 mm above the liquid. A vertical magnetic field interacted with the arc current at the molybdenum surface where the electrons came out radially. This impelled the arc spot around in the retrograde direction with a speed in these studies of about 120 meters per second. The racing spot, a horizontal line located on the molybdenum near the mercury was observed by a probe and oscillographic method and found to be fifteen mm long, more or less. It was most intense at the leading edge, falling off exponentially to the rear. An increase in arc current compressed the spot making the length less, the head more intense and the current density decrement along the length greater. A stronger magnetic field acted similarly. The speed limiting factor for the retrograde motion is evidently at the very front edge of the elongated racing spot. The arc current density here was about four hundred amperes per sq cm.

F7. Properties of a dc Arc in a Magnetic Field. L. P. WINSOR, *Rensselaer Polytechnic Institute,* AND T. H. LEE, *General Electric Company.*—An experimental study has been made of the motion of a direct current arc between plane parallel electrodes when subjected to a perpendicular magnetic field. Arc currents up to 100 amperes in fields up to 200 gauss have been investigated. In the lower current range, the cathode spot advances uniformly in the proper direction at the same rate as the column. The velocity increases nonlinearly with an increase in either the current or the field intensity, and is essentially independent of the electrode material. Except at the lowest currents, the anode spot tends to lag behind the column and advances in discrete steps in a similar manner to that observed by others for a moving anode. When the current is increased beyond certain limits, the cathode spot also tends to lag behind the column and advance in similar discrete steps. The average velocity is reduced appreciably from that observed for the first mode. The transition current is found to increase with an increase in field strength and to be dependent on the cathode material. The results have been correlated by photographs of the arc, the anode and cathode tracks, and cathode-ray oscillograms of the arc voltage. Anode current densities in the range between 10 000 and 100 000 amperes per square centimeter have been measured.

F8. Electric Field Measurements in Glow Discharges by Electron Beam Probe Techniques.* ROGER W. WARREN, *University of California, Berkeley.*—Electric field measurements have been made in the cathode fall region of glow discharges in He, Ar, N_2, H_2, and air. In most cases, the absolute values of the fields have been determined within a few percent. The measuring technique, a null method using an electron beam or a probe, permits high accuracy, high sensitivity, and continuous measurement and recording traversing a whole discharge within minutes. Fields in the positive column and fields in striations of a few volts/cm can also be faithfully recorded. Simultaneous recording of wall potentials and luminosity along the discharge as well as some spectroscopic studies have been made. The measurements indicate that at high pressures greater than mm Hg, the current carriers in the cathode fall region are entirely positive ions moving from the negative glow to the cathode with a velocity proportional to $(X/p)^{\frac{1}{2}}$. At lower pressures, complications occur.

* This work was supported at various stages by the Office of Naval Research.

F9. Very High Temperature Arcs. WOLFGANG FINKELNBURG, *Siemens-Schuckert Research Laboratories.*—Spectroscopic methods used for measuring temperatures in the range from 4000 to 50 000°K are discussed, and evidence is stated for the belief that, within the present limits of accuracy of measurement, thermal equilibrium exists between electrons, atoms, and ions in arcs at atmospheric pressure. As an example of high-temperature arcs, a 500-ampere arc in atmospheric pressure argon or nitrogen is described in which temperatures up to 30 000°K have been measured by G. Busz and the author. A detailed paper on this arc, which in addition to its high temperature and complete ionization in its axis is characterized by an operating voltage of as low as 8 volts, is to be published in volume 138 of the *Zeitschrift für Physik.*

G2. Wide-Range Electronic Tuning of Microwave Cavities in Gas Atmospheres. FRANK R. ARAMS AND HANS K. JENNY,

Radio Corporation of America.—Methods for electronically tuning microwave cavities using the principles of space-charge tuning and of spiral beam electronic tuning in the presence of a low-pressure gas are described. Very wide tuning ranges have been measured. One cavity was tuned over a frequency range from 3280 to 4350 Mc, and from 3280 to 2540 Mc, or ±30 percent. These values are compared to some measurements made in vacuum. Another was tuned over a range from 9170 to 10 800 Mc, or 18 percent. These values are compared to some measurements made in vacuum. A semiquantitative theory for electronic tuning in gas atmospheres is presented. Limitations of the method are given. Some microwave gas breakdown measurements are described.

G4. Influence of Space Charge on the Potential Distribution in Mass Spectrometer Ion Sources. W. M. BRUBAKER, *Consolidated Engineering Corporation.*—The role of space charge as a factor influencing the potentials and gradients in a mass spectrometer ion source of the electron bombardment type is calculated. Planar equipotential surfaces are assumed, and the analysis then becomes that of a plane parallel positive ion diode. The "cathode" of the ion diode is either emission-limited or space-charge-limited. The analysis considers the charge of the electrons in the ionizing sheet and the charge of the positive ions in the diode. At a critical gas pressure the influences of the positive and negative charges on the potential of the ionizing region are equal and opposite for small ionizing current. This concept leads to a pressure normalization in terms of the critical pressure. When this is done, sets of universal curves can be made for specific source geometries. Each set of curves gives the potentials and gradients in the source uniquely as a function of the normalized gas pressure and the ratio of the ionizing electron current density to the repeller voltage. A knowledge of the potential of the ionizing region is essential in appearance potential studies. Experimental data confirm the predictions of the theory.

G5. Studies of Active Nitrogen in Rare Gases. CARL KENTY, *General Electric Company.*—A Tesla spark between a flat W electrode and a W point 1½ mm distant in 0.1 mm N_2+200 mm A produces a blue flame having (a) an initial speed away from the flat electrode of 150 cm sec^{-1} resulting from rebounding neutralized ions, (b) a concentration of sputtered W atoms of 2×10^{11} cm^{-3}, (c) an ion concentration of $10^7-5\times10^9$ cm^{-3}, and (d) a concentration of N_2 and N metastables $>10^{12}$ cm^{-3}. The spectrum consists of low energy (mostly <3.7 V) W arc lines and the N_2 $2d$ positive bands, $C^3\pi_u(v_0=11\text{ V})\rightarrow B^3\pi_g$. With Ta electrodes, Ta 5212.8 (2.37 V) is strong showing that N(^{2}D) metastables (2.37 V) are present. $N_2(C^3\pi_u)$ may be excited by a collision of N(^{2}D) with $N_2(w)$ at about 8.6 V. The flame excites Na and (weakly) Hg but not Ba II. A 60~, 20 ma discharge in 0.1 mm N_2+200 mm He, Ne, A, or Kr produces a 5-sec orange afterglow (AG) due mainly to v_6(8.9 V) and v_4(8.2 V) of $N_2(B^3\pi_g)$, excited probably by the N_2 metastables w and/or a' through collisions (possible with slow electrons, following Nicholls). The AG excites Na, Hg, Cd, Zn, Tl, Mg, and Ba II. With Xe+N_2 no orange AG exists because the necessary carriers lose their energy to Xe (8.28 V); the Xe continua are emitted. Ba II 4550 (7.9 V) is strong showing a carrier is present of energy >7.9 V and <8.28 V, probably $N_2(a')$, placed tentatively by Gaydon at 8 V.

G6. Afterglow of Solid Nitrogen. HERBERT P. BROIDA, *National Bureau of Standards.*—Products from discharges through nitrogen have been condensed on surfaces at 4.2°K.[1] During condensation the solid material emits a yellow-green glow with occasional flashes of blue. After the supply of nitrogen is stopped and the discharge removed, a green glow continues for several minutes. Spectra of these glows have been obtained between 3100 and 9000 A with a Hilger E-2 spectrograph with glass optics and a Hilger $f/4$ spectrograph with quartz optics. These spectra show that the radiation is the same as that obtained from the bombardment by electrons of solid N_2.[2,3] It is possible that the glow is related to the formation of a solid lattice of molecular and atomic nitrogen. Evidence supporting this suggestion, obtained from discharges with other molecules, will be given. The preliminary experiments have been made with relatively simple apparatus in which nitrogen gas from commercial cylinders has been excited in a glass with an electrodeless discharge (2450 megacycles) at pressures of a few mm. Products from the discharge are pumped rapidly and continuously from the discharge region to a vessel immersed in liquid helium.

[1] H. P. Broida and J. R. Pellam, Phys. Rev. (to be published).
[2] L. Vegard, Nature 113, 716 (1924).
[3] McClennan, Ireton, and Thompson, Nature 118, 408 (1926).

CONTENTS—continued

Letters to the Editor

Measurement and Interpretation of Conductance of P-Type Inversion Layers on Germanium
G. A. deMars, H. Statz, and L. Davis, Jr. 539

Recombination of Surface States at the Germanium-Germanium Oxide Interface
H. Statz, L. Davis, Jr., and G. A. deMars 540

Spin-Lattice Relaxation Time of the Si^{29} Nucleus in Pure Fused Silica
G. R. Hoffman, John H. Anderson, and W. Koth 542

Superconductivity of Dilute Alloys
D. L. Dexter 543

Superconductivity of Dilute Magnesium Alloys
Bernard T. Salkovitz and A. I. Schindler 543

Properties of Noise Generating Functions
Peter G. Bergmann 544

Photoconductivity in PbS Photocells
E. S. Rittner and S. Fine 545

Some Cases of Ferroelectrics
A. N. Holden, B. T. Matthias, W. J. Merz, and J. P. Remeika 546

Surface-Controlled Electroluminescence and Light Amplification in Phosphor Films
D. A. Cusano 546

Theoretical Basis for Light-Amplifying Phosphors
Ferd E. Williams 547

Anomaly in the Low-Temperature Atomic Heat of Silver
P. H. Keesom and N. Pearlman 548

Absorption and Plasma Oscillation in Semiconductors with Magnetic Fields
Benjamin Lax and Laura M. Roth 548

Space Charge Layer Near the Surface of a Ferroelectric
Werner Känzig 549

Electromagnetic Resonance in Electronic Conductors
R. B. Dingle 550

Heat Capacity and Entropy of Liquid He^3 from 0.23 to 2°K: Nuclear Alignment in Liquid He^3
Bernard M. Abraham, Darrell W. Osborne, and Bernard Weinstock 551

Effect of Oxygen Gas on the Photoconductivity of BeO
M. Sakamoto, S. Kobayashi, and S. Ishii 552

X-Ray Analysis of Compressed KCl Powder Crystals
J. G. Kereiakes 553

Infrared Absorption of Liquid and Solid Hydrogen
Elizabeth J. Allin, W. F. J. Hare, and R. E. MacDonald 554

Dimensional Resonance in Lagrangian Quantum Mechanics
W. K. Burton 555

Sensitivity of a Nuclear Radiation Detector by Means of Electroluminescent Phosphors
W. Low 556

Quadrupole and Spin Resonance in Indium Antimonide
G. Dresselhaus, A. F. Kip, C. Kittel, and G. Wagoner 556

Proceedings of the American Physical Society

Minutes of the Seventh Annual Gaseous Electronics Conference (Held under the Joint Sponsorship of the Division of Electron Physics, American Physical Society, and the College of Engineering, New York University) at New York, New York, on October 14-16, 1954 557

CONTENTS

THE PHYSICAL REVIEW, SECOND SERIES

Partition Function for a System of Interacting Bose-Einstein Particles . . . . . T. D. Lee and C. N. Yang

Bose-Einstein Condensation of an Imperfect Gas . . . . . T. D. Lee and C. N. Yang

New Method for Measuring the α-Particle Range and Straggling in Matter . . . . . Connor Anderson

Mechanism of the Termination of the Geiger Plateau Region . . . . . and A. D. Krumbein

High-Frequency Gas Discharge Plasma in Hydrogen . . . . . David and Stanborough C.

Ambipolar Diffusion in a Magnetic Field . . . . . Alberti

Slope of the λ Curve of Liquid Helium . . . . . K. R. Atkins

Phenomenological Theory of Townsend Breakdown in Dielectrics . . . . . P. J.

Helium II Film Transport. I. The Role of Substrate . . . . . and Henry A.

Electron Spin Resonance Absorption in Metals. I. Experimental . . . . . and A.

Electron Spin Resonance Absorption in Metals. II. Theory of Electron Diffusion and the Skin Effect . . . . . Freeman J.

Oscillatory Thermomagnetic Properties of a Bismuth Single Crystal at Liquid Helium Temperatures . . . . . and J. Babiskin

Cyclotron Resonance of Electrons and Holes in Silicon and Germanium . . . . . G. Dresselhaus, A. F. Kip, and C.

Measurement of Shot Noise in CdS Crystals . . . . . C. I.

Interaction between Spin Waves and Conduction Electrons in Ferromagnetic Metals . . . . . Elihu Abrahams

Theory of Ionic Ordering, Crystal Distortion, and Magnetic Exchange Due to Conduction Electrons . . . . . J. H. Van Vleck and A. L.

Etch Patterns and Ferroelectric Domains in $BaTiO_3$ Single Crystals . . . . . John A. and Walter J.

Mobility of Slow Electrons in Polar Crystals . . . . . and David

Neutron Irradiation Effects in Cu and Al at 80°K . . . . . A. W. McReynolds, W. Augustyniak, Marilyn and D. B.

Paramagnetic Resonance Spectrum of Some Doubly Activated Fluorspars . . . . . W.

Temperature Dependence of the Width of the Band Gap in Several Semiconductors . . . . . Richard H.

Polarization Effects on the Brightness of Waves of Electroluminescence . . . . . Frank

Magnetic Rotation Phenomena in a Polycrystalline Ferrite. II . . . . . David

Direct Observation of Domain Rotation in Ferrites . . . . . and Charles L.

Axial Ratios in Hexagonal Crystals . . . . . J. W.

Properties of Germanium Doped with Nickel . . . . . W. and H. H.

Stopping Cross Sections of Metals for Protons of Energies from 400 to . . . . . D. Wayne Gregg, John and James C.

Cesium Inversion in CsF and RbF . . . . . G. Bennett, W. and H.

Equilibrium Charge Distribution of Energetic Nitrogen Ions . . . . . and

Orientation and Alignment of Sodium Atoms by Means of Polarized Resonance Radiation . . . . . William Bruce

Angular Correlation of Photons from Positron Annihilation in Light Metals . . . . . and A. T.

Magnetic Scattering of Slow Neutrons by Gaseous Oxygen . . . . . and R. M.

Electron Impact Broadening of Spectral Lines . . . . . and H.

Extension of Hund's Rule . . . . . G. F.

Electric Quadrupole Interaction and Spin Echoes in Crystals . . . . . and A.

Spin Echoes with Four Pulses—An Extension to n Pulses . . . . . and D.

Dirac Bracket Transformations in Phase Space . . . . . and Irwin

(Continued on inside cover)

Printed in U.S.A.